



Empiricité en linguistique et grammaire de Montague : la sémantique en 5P et la compositionnalité

Gabriel G. Bès

► To cite this version:

Gabriel G. Bès. Empiricité en linguistique et grammaire de Montague : la sémantique en 5P et la compositionnalité. [Rapport de recherche] Université Blaise-Pascal, Clermont-Ferrand. 2001. hal-01101344

HAL Id: hal-01101344

<https://hal.science/hal-01101344>

Submitted on 8 Jan 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Empiricité en linguistique et grammaire de Montague : la sémantique en 5P et la compositionnalité

avec en annexe

Grammaire de Montague et ambiguïté

Gabriel G. Bès

GRIL, Université Blaise-Pascal, Clermont-Ferrand

Rapport de recherche, GRIL, Université Blaise-Pascal, Clermont-Ferrand, 2001

Résumé

L'objectif de ce document est de spécifier les choix sur la sémantique dans le paradigme 5P et en particulier quels types de rapports sont souhaités entre syntaxe et sémantique. Ces choix sont étayés par une analyse détaillée de la grammaire de Montague et de la notion de compositionnalité en linguistique.

Le § 1 expose un ensemble hétérogène de difficultés rencontrées, ayant trait au flou terminologique et conceptuel. Le § 2 organise les apports de la grammaire de Montague et cible ceux qui seront discutés dans le document : idées qu'il n'y a pas de différence entre langues naturelles et langages artificiels des logiciens, que l'intérêt principal de la syntaxe est de conduire à la sémantique et que l'objectif de la sémantique est de caractériser les notions d'énoncé vrai et de conséquence. Le § 3 pose la question de l'empiricité de ces trois idées et montre qu'aucune ne peut être vue comme une hypothèse d'une science du réel, puis interroge la notion de compositionnalité dans ses variantes générales et dans le cadre stricte de UG. Dans le § 4, l'auteur reformule le système de spécification d'un fragment de l'anglais dans PTQ dans le modèle des grammaires formelles et observe que la grammaire résultante, d'une part, incorpore dans son vocabulaire terminal des expressions qui ne sont pas de l'anglais, d'autre part, a parmi ses règles des règles de type 0. Il expose quelques inadéquations descriptives : sur- et sous-génération de la grammaire et erreurs d'interprétations produites par les fonctions de traduction. Dans l'alinéa suivant (§ 5), Bès montre comment les exemples non générés peuvent l'être en développant un peu la grammaire de PTQ, dans le strict respect des restrictions du système, mais en utilisant ses permissivités, au premier rang desquelles effacement et permutation. Le § 6 fait le bilan et réévalue la grammaire de Montague et les interprétations qui en sont faites pour aboutir à la conclusion que l'intégration de la syntaxe et la sémantique des langues naturelles et des langages artificiels des logiciens dans une unique théorie mathématique n'a jamais vu le jour jusqu'ici. Le § 7 fait ressortir différents points qui différencient les langues naturelles des langages artificiels des logiciens, caractérisant ainsi la « naturalité des langues naturelles ». Enfin, le § 8 présente un programme de recherche en sémantique pour le paradigme 5P : y est exposé d'une part le choix d'une compositionnalité comprise comme la systématisme de la traduction effectuée par les Fonctions sémantiques et non comme un homomorphisme entre structures algébriques de la langue naturelle et l'univers de dénotation, d'autre part un plaidoyer pour la définition d'un système d'observation systématique susceptible d'exprimer les observations d'ordre sémantique dans un langage ad hoc

L'annexe au document principal étudie le traitement de l'ambiguïté dans la grammaire de Montague à travers deux questions : la possibilité formelle de l'ambiguïté d'assignation catégorielle dans le système de Montague et les répercussions systémiques de cette ambiguïté.

Gabriel G. Bès

**Empiricité en linguistique
et grammaire de Montague :
la sémantique en 5P et la compositionnalité**

avec en annexe

Grammaire de Montague et ambiguïté

**GRIL
Université Blaise-Pascal
Avril 2001**

Travaux non définitifs

Gabriel G. Bès
Université Blaise-Pascal/GRIL
34 avenue Carnot • F-63037 Clermont-Ferrand cedex
Gabriel.Bes@univ-bpclermont.fr

Pages liminaires / Preliminaries

Préface / Foreword	III / V
---------------------------------	----------------

Empiricité en linguistique et Grammaire de Montague : la sémantique en 5P et la compositionnalité. Résumé / Empiricism in linguistics and Montague Grammar: Semantics in 5P and Compositionality. Abstract	VII / XI
---	-----------------

Annexe/ Annexe

Grammaire de Montague et ambiguïté. Résumé / Montague Grammar and ambiguity. Abstract	XV / XVII
--	------------------

Préface

Quelqu'un a dit qu'après Kant, celui qui souhaite s'occuper de métaphysique doit être soit pro-Kant, soit anti-Kant, mais qu'il ne peut pas être a-Kant.

On doit dire quelque chose d'analogue par rapport à Montague : si l'on souhaite s'occuper de la sémantique des langues naturelles, il est impossible de ne pas prendre position par rapport à lui.

C'est ce que ce document prétend faire avec, comme point de départ, le parti-pris méthodologique basique du Paradigme 5P (désormais 5P), qui consiste à essayer de pratiquer la linguistique comme science empirique. Le lecteur intéressé trouvera un résumé très succinct de 5P dans la Section finale (§ 8) du premier travail ci-après, mais il doit savoir que la compréhension de l'analyse de Montague ne nécessite nullement la connaissance du formalisme de 5P.

L'optique choisie n'épuise pas la richesse des apports de Richard Montague. Celui-ci est essentiellement revisité dans ce document en relation à sa notion de compositionnalité. Cela conduit à étudier sa position sur la syntaxe des langues naturelles, sur la traduction de la syntaxe de la langue naturelle en logique intensionnelle et sur son interprétation dans un univers de dénotation.

En faisant cette re-visite, nous avons été amené à revisiter aussi les idées fondatrices de la démarche de Montague et, en particulier, son idée selon laquelle il n'y aurait pas de différences significatives entre les langues naturelles et les langages artificiels des logiciens.

Après analyse de « Universal Grammar » (UG) et de « The Proper Treatment of Quantification in Ordinary English » (PTQ), nous avons décidé de ne pas suivre Montague sur des points importants de son projet. Non seulement sa syntaxe, soi-disant formelle, est quasiment inexistante – point qui ne devrait pas susciter beaucoup de controverses –, mais son exigence de compositionnalité, telle qu'il la comprend, nous apparaît comme non motivée dans le cadre d'une linguistique empirique. Ce point, en revanche, devrait susciter plus de résistances.

Son idée sur la compositionnalité a permis, certes, d'avancer dans la possibilité d'associer les expressions d'une langue naturelle à des formules d'une logique quelconque. Mais elle a été formulée dans un cadre général visant à assimiler langues naturelles et « langages artificiels des logiciens », pour reprendre sa célèbre formule. De ce fait, pour forcer la compositionnalité à laquelle il prétend, il a dû détourner le regard des aspects de la langue naturelle qui la distinguent des langages artificiels, qui se manifestent en particulier dans sa syntaxe, qui la rendent intraitable avec les formalismes des langages artificiels des logiciens et qui, en les explicitant, définissent ce que nous avons caractérisé comme *la naturalité de la langue naturelle*.

Dans l'analyse selon l'angle qui a été le nôtre de PTQ et de UG, nous avons été confronté à des points qui nous avaient échappé auparavant, et qui, sauf mauvaise information de notre part, avaient échappé aussi à ceux qui ont présenté Montague. On pense, en particulier, à la question, toujours présente en LN, de l'assignation catégorielle ambiguë des expressions linguistiques, élémentaires ou non. Notre analyse nous amène à la conclusion que si l'observation de ce phénomène existe dans les langues naturelles, son expression formelle dans le système de Montague pose de graves problèmes : pour en rendre compte dans son système d'ensemble, qui

incorpore la traduction ou l'interprétation, il faut utiliser les capacités formelles démesurées de sa syntaxe, ce qui rend celle-ci non calculable.

En essayant de traiter des données d'observation de la langue naturelle, comme par exemple celui de l'assignation catégorielle ambiguë des expressions ou de phénomènes de linéarité, nous avons découvert, de fil en aiguille, tout ce qui n'est pas dit dans le système de Montague et qui laisse tant de portes ouvertes pour faire ce que l'on veut, malgré les apparences de rigueur et de contraintes. Et ces portes doivent être laissées impérativement ouvertes pour pouvoir assimiler langues naturelles et langages artificiels des logiciens. Dans l'étude des langues naturelles en tant que langages-objet, cette idée directrice est, nous semble-t-il, le plus fort et le plus important des antécédents d'importation de formalismes et de manières de voir qui ont été développés avec succès ailleurs, pour traiter d'autres objets avec d'autres fonctionnalités. Et dont la justification pour connaître la langue naturelle – étudiée empiriquement comme langage objet – est loin d'être faite.

Nous sommes conscients que notre problématique de fond – quelle sémantique pour 5P – devrait intéresser un public extraordinairement restreint. Comme les choix auxquels nous avons abouti pour positionner 5P nous semblent raisonnés et comme ce raisonnement transite par l'évaluation des points centraux de la grammaire de Montague, nous avons pensé qu'il était utile de faire circuler les deux travaux ci-après auprès de ceux qui se sont intéressés à Montague ou qui sont susceptibles de le faire. Le premier de ces deux travaux, « Empiricité en linguistique et Grammaire de Montague : la sémantique en 5P et la compositionnalité », définit le cadre de l'analyse, distingue dans ce cadre UG de PTQ et, en fondant les choix faits par rapport à 5P, esquisse les caractéristiques susceptibles de définir la naturalité de la langue naturelle. Le deuxième de ces deux travaux, « Grammaire de Montague et ambiguïté », étaye des points du raisonnement du document précédent se centrant sur la problématique formelle qui surgit dans le système de Montague à partir de l'assignation ambiguë des expressions linguistiques à des catégories différentes.

Ce document, dont nous assumons toute la responsabilité, a bénéficié, pour sa version finale, de longues discussions avec Luisa Coheur. Le document, déjà prêt dans sa version finale, hormis quelques détails, à la fin décembre 2000, n'est pas définitif. Nous le faisons circuler dès maintenant, car nous serions heureux et reconnaissant de recevoir des critiques, commentaires, discussions et, plus généralement, tout type de réactions, afin de pouvoir lui donner une forme plus définitive, et nous les remercions par avance.

À Clermont-Ferrand, le 12 avril 2001

Foreword

Someone once said that, after Kant, he who wishes to study metaphysics must either be in favor of the German philosopher, or against him, but he cannot ignore him.

Something similar can be said concerning Montague : he who wishes to consider natural language semantics must take a stance on him.

That is what this document is intended to do, guided by the basic methodological choice of the 5P Paradigm (from hereafter *5P*), which consists of trying to do the practice of linguistics as an empirical science. The reader will find a brief "résumé" of 5P at the last Section (§ 8) of the first paper attached below, but he must know that the understanding of the analysis of Montague's work does not require to be aware of the 5P formalism.

The chosen perspective does not begin to exhaust the richness of Richard Montague's work. This is primarily revisited with respect to the notion of compositionality, which leads to the study of his views on natural language syntax, on translation of natural language syntax into intensional logic and its interpretation in a universe of denotations.

Whilst studying Montague, we have found it necessary to reappraise fundamental Montague ideas and in particular, his idea that there are no significant differences between natural languages and the artificial languages of logicians.

After analysis of "Universal Grammar" (UG) and "The Proper Treatment of Quantification in Ordinary English" (PTQ), we decided not to follow Montague on several important points of his project. It is not just that his supposedly formal syntax is nearly non existent (rather a non-controversial point), but that his requirements on compositionality, as he conceived it, do not appear motivated in the perspective of empirical linguistics. On the other hand, this point is apparently more controversial.

His idea on compositionality certainly allowed progress on the issue of associating natural language expressions to formulae of any logic. However, it was formulated in a general manner in order to assimilate natural languages with the "artificial languages of logicians", to quote his celebrated phrase. By forcing the compositionality which he was trying to obtain, he was obliged to ignore aspects of natural languages which distinguish them from artificial languages. This is particularly shown in his syntax which makes it unworkable with the standard formalisms of logicians' artificial languages, defining what we have characterized as the *naturalness of natural languages*.

Whilst carrying out PTQ and UG analysis, from our particular point of view, we were confronted by aspects of Montague's work which had previously escaped our attention and, unless we are mistaken, had also escaped the attention of those who had previously analyzed Montague's work. Specifically, the question, which is always present in natural languages, of the phenomenon of categorial ambiguity of linguistic expressions, elementary or not. Our analysis of PTQ and UG brought us to the conclusion that the formal expression of this phenomenon, which, empirically, certainly does exist in natural languages, poses serious problems to Montague's system: if we want to account for it in his general system, incorporating translation and interpretation, it is necessary to make use of the unlimited power of his syntax, which would make it non-computable.

In trying to account for observational phenomena of natural languages, as, for example, categorial ambiguity of expressions or linearity phenomena, we discovered many things as we proceeded, which are not mentioned in Montague's system and which leave so many doors open to do what we want, despite the apparent rigor and constraints. And these doors must imperatively be left open in order to assimilate natural languages to logicians' artificial languages. Throughout the history of the study of natural languages as object languages, this guiding idea of trying to assimilate natural languages to artificial ones seems to us to be one of the most important antecedents to the attitude of importing formalisms. Concepts elaborated with success in other fields, but in order to treat other objects with other functionalities than those of natural languages, were introduced into the study of these, whilst neglecting a rigorous description of the justifications for the importations, justifications that are necessary if one hopes to understand natural languages empirically studied.

We are aware of the fact that the basic problem (What semantics for 5P?) can only be of interest to very few. Because the choices we have reached in order to position 5P appear rationally founded to us, and because this reasoning arises from the evaluation of the central points of Montague Grammar, we thought that it would be useful to circulate this document with the two attached papers amongst those who were interested in Montague or who might become so. The first of these two papers, "Empiricity in Linguistics and Montague Grammar: the Semantics in 5P and Compositionality" defines the analytical framework, distinguishes in terms of this framework UG from PTQ and, while supporting choices on 5P, sketches characteristic features that are likely to define the naturalness of natural languages. The second paper, "Montague Grammar and Ambiguity", reinforces the reasoning of the previous one. It is centered on the problematic of formalisms which arise in the Montague system from an ambiguous assignment of linguistic expressions to different categories.

This document, for which we assume full responsibility, has benefited from long discussions with Luisa Coheur. The document, ready in its final version at the end of December 2000, minor details excepted, is not a definitive one. We make it circulate from now because we'll be happy to receive any criticisms, comments, discussions and, more generally, any reactions whatsoever, in order to give it a more solid form and we'll be very grateful for them.

Clermont-Fnd, April 12th 2001.

Empiricité en linguistique et Grammaire de Montague : la sémantique en 5P et la compositionnalité

Résumé

L'objectif essentiel du document étant de spécifier les choix sur la sémantique en 5P, et, en particulier, quel type de rapports on y souhaite entre syntaxe et sémantique, Montague a été revisité en ce qui concerne la syntaxe qu'il vise pour les langues naturelles et les rapports de celle-ci avec la traduction en LI (Logique Intensionnelle) et son interprétation dans un univers de dénotation (§1).

Les difficultés rencontrées (§1.1) ont été grandes. Elles viennent très généralement du très grand flou terminologique prédominant pour traiter du cadre méthodologique de la linguistique : des termes comme *hypothèse*, *théorie*, *formel*, *langage*, *catégorie* admettent des interprétations très peu rigoureuses, de même que *compositionnalité*, *Grammaire de Montague* et *principe frégéen*. Les choix terminologiques sont fixés et, parmi d'autres, *Grammaire de Montague* est strictement réservée à UG (« Universal Grammar ») et à PTQ (« The Proper Treatment of Quantification in Ordinary English »). Elle est analysée selon les préoccupations d'une science de l'empirie. Les apports de la Grammaire de Montague sont organisés (§ 2) autour des trois pôles suivants :

- i Idées générales et profondes qui se situent à la lisière de la logique, de la linguistique et de la philosophie du langage.
- ii Spécification syntaxique des LN (Langue(s) Naturelle(s)) et traduction des suites de la LN en formules LI.
- iii Définition d'une logique intensionnelle (LI) avec son univers de dénotation et ses règles de calculabilité.

Le pôle (iii) n'est en rien abordé. Le pôle (i) se laisse à son tour synthétiser par les trois Idées suivantes (citations textuelles de [Chambreuil & Pariente 90]) :

- Idée 1 Il n'y a pas de différences théoriques fondamentales entre les langages naturels et les langages artificiels des logiciens ; la syntaxe et la sémantique de ces différents langages relèvent d'une même théorie mathématique (p. 19).
- Idée 2 L'intérêt principal de la syntaxe est de conduire à la sémantique (p. 64).
- Idée 3 L'objectif principal de la sémantique est de caractériser les notions d'énoncé vrai et de conséquence (p. 64).

Les interrogations du travail (§3) portent sur l'empiricité des Idées 1 à 3, et sur l'exigence de compositionnalité en tant qu'homomorphisme algébrique, dans laquelle s'instancie l'Idée 1. L'analyse est conduite séparément par rapport à UG et à PTQ (§3.2). On aboutira à la conclusion qu'il n'y a rien dans la Grammaire de Montague qui autorise à considérer l'Idée 1 comme une hypothèse d'une science empirique. En ce qui concerne les Idées 2 et 3, elles sont considérées comme le résultat d'un choix, qui peut être différent, et non comme un impératif absolu pour étudier la sémantique des langues naturelles.

Cinq notions différentes de compositionnalité sont distinguées (§3.2). Les définitions de UG qui fixent celle d'homomorphisme unique d'assignation d'interprétation ou de traduction sont rappelées (§3.2.1). De cette présentation, résultent trois Observations.

Ob. 1 Montague n'a pas utilisé *compositionnalité* ou ses dérivés comme des notions techniques et on voit mal qu'il attribue ce « principe » à Frege. Par son expression « Fregean interpretation » on peut comprendre qu'avec ses notions à lui d'homomorphisme algébrique et sa manière à lui d'organiser l'univers d'interprétation, il peut rendre compte des observations de Frege sur *Sinn* et *Bedeutung*.

Ob. 2 Montague n'introduit aucune restriction sur aucun des éléments du quintuplet qui définissent formellement sa notion de « langage », à l'exception de celles qui devraient garantir la non-ambiguïté, et aucune restriction sur la relation *R* qui devrait associer un langage éventuellement ambigu à un langage non ambigu.

Ob. 3 Étant donné les définitions de UG, il est impossible, dans la description des LN, d'assigner une expression élémentaire d'un « langage non ambigu » de manière ambiguë à des ensembles indexés différemment afin obtenir des assignations de significations différentes, car, si on le faisait, on obtiendrait des résultats inadéquats. Ce point est repris dans l'Annexe *Grammaire de Montague et ambiguïté* ; cf. ci-dessous son résumé.

L'analyse de PTQ, en ce qui concerne la syntaxe de la langue naturelle et sa traduction en formules de LI (§4), est faite inductivement, en fonction de la présentation concrète de PTQ, et non dans le cadre d'un formalisme général, inexistant. Le système PTQ est noté par *pSLF* car en fait c'est un pseudo-formalisme spécifiant un langage ambigu, et cela, tout particulièrement, par le recours à des « manipulations » qui ont l'aspect de règles et qui sont des descriptions en anglais de procédures supposées chercher des éléments dans des listes, les substituer par d'autres, et qui utilisent des opérations d'effacement et de substitution qui n'obéissent à aucune contrainte.

Pour essayer de comprendre le pouvoir expressif de PTQ, les règles de *pSLF* ont été reformulées (§4.1.3) par des règles de production de la hiérarchie chomskyenne, et, lorsque c'est possible, en utilisant l'application fonctionnelle des grammaires catégorielles, et les entités utilisées comme expressions élémentaires sont distribuées dans une partition comprenant 5 ensembles différents (§4.1.4).

Les permissivités utilisées dans *pSLF* (§4.1.5) en font un système irrestreint qui, par ailleurs, ne fixe aucune limite détectable entre les objets qu'il spécifie et ceux qui peuvent résulter d'une observation directe des expressions de la langue naturelle. Une de ses rares limitations, même si elle n'est pas formellement exprimée, est celle de ne pas pouvoir attribuer une même expression élémentaire à plus d'une catégorie.

Le rappel est fait (§4.2) des éléments de LI et de la traduction en LI de *pSLF*. L'analyse d'exemples de PTQ (§4.3) ne vise pas à détecter, en se donnant l'anglais comme domaine d'observation, toutes les absences ou erreurs descriptives de *pSLF*. Il s'agit plutôt d'attirer l'attention sur trois exemples qui illustrent trois absences caractérisées du domaine traité dans PTQ. Ce sont :

- Exemple 1 Syntagmes prépositionnels comme des modificateurs des noms
- Exemple 2 Lemmes verbaux avec plus d'une sous-catégorisation
- Exemple 3 Emplacement des adverbes

Ces absences n'attestent pas de l'incapacité formelle de PTQ à les traiter. C'est plutôt le contraire (§5) : en utilisant les permissivités syntaxiques et de traduction de PTQ, on peut traiter ces cas, tout en respectant les contraintes de la traduction, et, en particulier, la traduction exigée des catégories syntaxiques en types de LI.

Les solutions utilisées sont parfaitement généralisables (§5.3). En utilisant les permissivités de pSLF et de la traduction, on peut trouver des solutions à beaucoup de problèmes plus généraux concernant les relations d'ordre et l'assignation ambiguë des expressions élémentaires à des catégories différentes. Mais le prix à payer est très fort : comme dans pSLF, il faut en syntaxe utiliser des symboles vides, effacer, permuter..., et faire des manipulations.

Des points importants de la Grammaire de Montague sont réévalués (§6). La compositionnalité, telle qu'elle est parfois présentée intuitivement, ne donne aucunement une idée vraie des enjeux théoriques et conceptuels de fond, et de la compositionnalité stricto sensu, telle que définie dans UG. La compositionnalité intuitive induit en plus des interprétations abusives sur la possibilité d'utiliser la compositionnalité stricto sensu dans le traitement automatique. Il n'est pas vrai non plus que Montague ait utilisé une version modifiée des grammaires catégorielles et qu'il ait introduit des contraintes sévères sur la syntaxe et la sémantique des langues naturelles en termes de sa notion de compositionnalité. Des points formels cruciaux pour pouvoir démontrer les effets de l'interprétation induite n'ont jamais été démontrés. Enfin, l'intégration de la syntaxe et la sémantique des deux types de langages – les langues naturelles et les langages artificiels des logiciens – dans une unique théorie mathématique, naturelle et précise, n'a jamais vu le jour jusqu'ici.

La compositionnalité stricto sensu visée par Montague n'est donc pas assimilable à une hypothèse d'une science empirique. Elle ne peut subsister que comme un commandement méthodologique – *l'homomorphisme syntaxe-sémantique tu construiras* – qui ne peut satisfaire qu' à ceux qui y croient.

Cependant, l'Idée 1 de Montague peut être utilisée comme un moyen indirect pour apporter des connaissances empiriques sur les langues naturelles (§7). Les morphismes algébriques sont utilisés dans les sciences empiriques, comme par exemple la cristallographie (§7.1), pour exprimer des connaissances testables sur le réel-observé. Leur utilisation n'est pas susceptible d'être importée telle quelle en linguistique pour aborder les problèmes traités, mais la compositionnalité de Montague et son Idée 1 à elle associée peuvent être ré-utilisées pour caractériser les langues naturelles et les différencier des langages artificiels des logiciens.

Si l'on a la trilogie :

- i Alg-L où algèbre décrivant syntaxiquement les expressions d'un langage L ;
- ii Alg-U, où algèbre décrivant l'univers de dénotation ;
- iii homomorphisme-u, associant Alg-L et Alg-U.

nous avons une situation où (i) va varier en fonction des variations de (ii) et ceci si l'on souhaite respecter (iii). Supposons que l'on ait deux langages L1 et L2. Si l'on fixe (ii) et que l'on se donne (iii) comme exigence, aussi bien pour $\langle \text{Alg-L1}, \text{Alg-U} \rangle$ que pour $\langle \text{Alg-L2}, \text{Alg-U} \rangle$, les différences entre L1 et L2 peuvent s'exprimer en caractérisant les différences entre Alg-L1 et Alg-L2, les deux étant requises pour satisfaire (iii) sur un même objet ou sur une même classe d'objets.

Ce schéma est celui des définitions par *genus* et *differentia*. Les algèbres et le respect de la compositionnalité stricto sensu de Montague nous donnent le *genus* alors que la comparaison des algèbres nous donne la *differentia*. Les caractéristiques suivantes, reprises dans les formulations I à IX (§7.2), spécifient les points par rapport auxquels on peut distinguer langues naturelles et langages artificiels des logiciens : (i) opération de permutation, (ii) opération d'effacement, (iii) caractérisation de R, (iv) utilisation de parenthèses indexées, (v) computabilité, (vi) assignation graphique des expressions élémentaires à des catégories différentes, (vii) existence possible d'ambiguïté dans les assignations catégorielles des expressions, (viii) utilisation nécessaire des postulats de signification, (ix) exigence d'une algèbre dérivée comme traduction des expressions de la langue naturelle. Les caractéristiques qui résultent de l'explicitation de ces neuf points définissent la *naturalité de la langue naturelle*.

Le choix pour 5P (§ 8) a été de ne se pas se plier au commandement méthodologique de la compositionnalité, et cela pour deux raisons : la première parce que la compositionnalité, n'étant pas empiriquement motivée, n'est pas souhaitée, la seconde parce qu'elle ne peut pas être obtenue dans 5P, si l'on devait se soumettre à l'« épreuve de compositionnalité » de [Janssen 97]. En effet, [Janssen 97] pour pouvoir aboutir à la compositionnalité qu'il souhaite, ne traduit pas une structure *s* dans une autre, mais il prend en entrée le processus qui permet l'obtention de *s*. Le domaine de la fonction de traduction est le processus représenté par la suite dérivationnelle, qui est soigneusement distingué du graphe final que ce processus dérivationnel permet de spécifier. Or l'utilisation du processus dérivationnel ré-introduit ce que, dans 5P, on veut soigneusement laisser dehors, à savoir des aspects algorithmiques dans des descriptions qui se veulent purement déclaratives. Les P2 de 5P ne décrivent pas des processus dérivationnels. Elles vont spécifier des modèles, dont les plus achevés sont des modèles avec un fléchage. On n'en récupère pas de dérivations ; elles n'existent pas ; par conséquent elles ne sont l'entrée de rien. Si l'obtention de l'homomorphisme algébrique souhaité en a besoin, la conclusion en est claire : celui-ci ne peut pas être obtenu.

Notes complémentaires

La **Note 1**, en analysant [Cooper 83] (utilisation du *Cooper storage* pour exprimer des différences de portée) et [Partee 79a] (utilisation d'une composante morphologique et des règles un peu plus contraintes en syntaxe), suggère une méthode pour vérifier les affirmations de §7.2 in fine, qui synthétisent la naturalité des langues naturelles : les développements, extensions, améliorations, etc. de pSLF ont été obligés de se donner, explicitement ou pas, des exigences de compositionnalité qui ne sont pas celles de Montague dans UG.

La **Note 2** illustre a contrario, à partir de [Cooper 83], un raisonnement non empirique sur la sémantique des langues naturelles et à partir de [Partee & al 90 chap 14] et de [van Benthem 86 1ère Partie] un autre qui est, cette fois, authentiquement empirique et qui porte sur la monotonie des différents déterminats et leur combinatoire possible avec la négation.

Empiricism in linguistics and Montague Grammar: Semantics in 5P and Compositionality

Abstract

The essential goal of this document is to specify the choices on semantics in 5P, and in particular, the desired relations of syntax and semantics. As so, Montague is revisited with respect to the way he deals with the syntax of natural languages, with the relations between natural language syntax and its translation into Intensional Logic (IL) and with its interpretation in a denotational universe (§1).

It was not an easy task (§1.1). Many difficulties come essentially from the fuzziness of terminology that prevails in the methodological vocabulary of linguistics : terms like *hypothesis*, *theory*, *formal*, *category* admit non-rigorous interpretations, as well as *compositionality*, *Montague Grammar* and *Frege's Principle*. The methodological choices are fixed and, besides others, *Montague Grammar* is strictly reserved for UG ("Universal Grammar") and PTQ ("The Proper Treatment of Quantification in Ordinary English"). Montague Grammar is analysed from the goals of an empirical science. Its contributions are organised around the three following centres :

- i General and deep ideas in the borderline of logic, linguistics and philosophy of language.
- ii Syntax specification of NL (Natural Language(s)) and translation of NL sequences into IL formula.
- iii Definition of IL with its denotational universe and its inference rules.

In this document, (iii) is not tackled. (i) can be synthetised by the following three Ideas (translation of textual citations from [Chambreuil & Pariente 90]) :

- Idea 1 There are no fundamental theoretical differences between natural languages and artificial languages of logicians; syntax and semantics of these different languages come under the same mathematical theory (p. 19).
- Idea 2 The principal goal of syntax is to lead to semantics (p. 64).
- Idea 3 The principal goal of semantics is to characterise the notions of true sentence and of entailment (p. 64).

The targeted questions (§3) of the document concern the empiricity of Ideas 1 to 3 and the requirement of compositionality conceived as an algebraic homomorphism, requirement in which Idea 1 is instantiated. Two separate analyses are made, the one on UG, the other on PTQ (§ 3.2). The conclusion is that there is nothing in Montague Grammar allowing to consider Idea 1 as a hypothesis of an empirical science. Ideas 2 and 3 are considered to be the result of a choice, that could be different, and not as an absolute requirement in the study of NL semantics.

Five different notions of compositionality are distinguished (§ 3.2). UG definitions on unique homomorphism of meaning assignement and of interpretation are reminded (§ 3.2.1). From this presentation, three Observations resulted.

Ob. 1 Montague has not made use of *compositionality* or its derivations as technical notions and it is strongly dubious that he assigned this “principle” to Frege. The text with his expression “Fregean interpretation”, contextually interpreted, can be understood as claiming that he, Montague, is able to account for Frege's observations on *Sinn* and *Bedeutung* by his own notions of algebraic homomorphism and his own way of organising the denotational universe.

Ob. 2 Montague does not introduce any restrictions on any of the elements of the 5th-tuple that formally defines his notion of “language”, with the exception of those that should guarantee the non-ambiguity. Futhermore, he does not introduce any restriction on the relation *R* that should associate a possible ambiguous language to a non-ambiguous one.

Ob. 3 Given UG definitions, it is not possible, in NL descriptions, to associate a basic expression of a “disambiguated language” in an ambiguous way to differently indexed sets in order to obtain different meanings ; inadequate results are obtained doing this. This point is analysed in more detail in the Annexe “Montague Grammar and ambiguity”; see the abstract below.

The analysis of PTQ inasmuch as it concerns NL syntax and its translation into IL formula (§ 4), is done inductively, in terms of the actual presentation, and not in terms of a general formalism, which is missing. The PTQ system is noted by *pSLF* as it is a pseudo-formalism specifying an ambiguous language. And this, particularly, by means of “manipulations” with the aspect of rules – i.e. English descriptions of procedures intended to search and to substitute elements in lists, and which use operations of deletion and/or substitution constrained by nothing.

In order to try to understand PTQ expressive power, *pSLF* rules were re-formulated (§ 4.1.3) by production rules of the Chomsky hierarchy, and, when possible, by functional application of categorial grammars. The entities used as basic expressions are distributed in a partition of 5 different sets (§ 4.1.4).

The lack of constraints in *pSLF* (§ 4.1.5) makes of it an unrestricted system which, furthermore, does not impose any explicit boundary between the objects it specifies and the ones that can be directly observed in NL expressions. One of its rare limits, though not formally imposed, is the impossibility of assigning one basic expression to more than one category.

The basic features of LI and of the translation of LI into *pSLF* are reminded (§ 4.2). The goal of the analysis of PTQ examples (§ 4.3), given English as the observational domain, is not to find *pSLF* descriptive errors or *pSLF* descriptive gaps but, rather, to take notice of three examples typically missing in the domain targeted by PTQ. These are :

- Example 1 Prepositional phrases as noun modifiers.
- Example 2 Verbal lemmas with more than one sub-categorisation.
- Example 3 Adverb positions.

These absences do not attest of a PTQ intrinsic incapacity of accounting for them. In fact, it is the opposite (§ 5) : thanks to PTQ permissiveness on syntax and translation, it is possible to account for them while satisfying the translation constraints and, in particular, the constraints on the translation of syntactic categories to IL types.

The proposed solutions can be easily generalized (§ 5.3). Thanks to PTQ permissiveness on syntax and translation, it is possible to account for many general problems concerning order relations and ambiguous assignment of basic expressions to different categories. But the price to pay is very high : as in pSLF, it is necessary in syntax to use empty symbols, to delete, to permute, ... , to make use of “manipulations”.

Important points of Montague Grammar are re-evaluated (§ 6). Intuitive compositionality, as it is sometimes presented, does not give a true idea neither of more deep theoretical and conceptual issues, nor of compositionnality *stricto sensu* as defined in UG. Furthermore, intuitive compositionality gives rise to abusive interpretations on the possibility of using *stricto sensu* compositionality in NL computing. Neither is it true that Montague used a modified version of categorial grammars and that his notion of compositionality introduces severe constraints on NL syntax and semantics. Crucial formal points needed for formal proving of induced interpretation were never demonstrated. Lastly, the integration of syntax and semantics of the two types of languages – natural languages and the artificial languages of logicians – in a unique mathematical theory, natural and precise, has never seen light.

Stricto sensu compositionality, such as it was aimed at by Montague, cannot be assimilated to a hypothesis of an empirical science. It can only survive as a methodological commandment – *syntax-semantics homomorphism you will build* – that satisfies only those who believe in it.

However, Montague's Idea 1 can be utilised as an indirect method of contributing to empirical knowledge on natural languages (§ 7). Algebraic morphisms are utilised in empirical sciences, e.g. crystallography (§ 7.1), in order to express testable knowledge on observed reality. The way used there cannot be imported into linguistics, but Montague compositionality and its associated Idea 1 can be re-used to characterise natural languages and distinguish them from artificial languages of logicians.

If we have the trilogy

- i Alg-L, algebra that describes syntactically the expressions of a language L;
- ii Alg-U, algebra describing the denotational universe;
- iii homomorphism-u, from Alg-L to Alg-U;

we have a situation where (i) is going to vary in function of variations of (ii) and this if it is wanted to respect (iii). Suppose that we have two languages L1 and L2. If (ii) is fixed and (iii) is assumed as a requirement on <Alg-L1, Alg-U> and <Alg-L2, Alg-U>, the differences between L1 and L2 can be expressed by the characterisation of the differences between Alg-L1 and Alg-L2, both satisfying (iii) on the same object or on a same class of objects.

This schema is the one of the definitions by *genus* and *differentia*. The involved algebras and the respect of Montague's compositionality, give the *genus*. The comparison of the algebras gives the *differentia*. The following characteristics, abstracted from formulations I to IX (§ 7.2), specify points discriminating natural languages and artificial languages of logicians : (i) permutation operation; (ii) deletion operation; (iii) characterisation of R; (iv) use of indexed parentheses; (v) computability; (vi) assignement of basic expressions to categories in terms of graphic characteristics; (vii) possible ambiguity in the assignement of expressions to categories; (viii) need of meaning postulates; (ix) need of a derived algebra as translation of NL expressions. Characteristics resulting from the explicitation of these nine points define the *naturalness of natural languages*.

The choice made in 5P (§ 8) is to not blindly obey the methodological commandment of compositionality, and this for two reasons : because compositionality, not being empirically motivated, is not wanted, because it cannot be obtained in the 5P pattern, should the “compositional requirements” of [Janssen 97] be satisfied. As a matter of fact, [Janssen 97], in order to obtain the compositionality he wants to have, does not translate a structure *s* into another one, but uses as input the process that allows obtaining *s*. The translating function domain is the process represented by the derivational sequence, which is carefully distinguished from the final graph that this derivational process allows to specify. In fact, the use of the derivational process re-introduces exactly what is wanted to avoid in 5P, i.e. algorithmic aspects in descriptions intended to be purely declarative. P2 (Properties) of 5P do not describe derivational processes. They specify models, the most complete of which are arrowing models. Derivational processes are not recovered; they do not exist and by consequence, they cannot be the entry of anything. If the obtention of the algebraic homomorphism wished by [Janssen 97] needs them, the conclusion is clear : the homomorphism cannot be obtained.

Complementary notes

Note 1 analyses [Cooper 83] (use of Cooper storage to express scope differences) and [Partee 79a] (use of a morphological component and of more constrained syntactic rules) and suggests a method to verify the final statements in § 7.2, which synthetise naturalness of natural language : pSLF developments, extensions, improvements, etc. explicitly or not, force requirements on composition which are not that of Montague in UG.

Note 2 illustrates *a contrario*, bearing on [Cooper 83], a non empirical argument on NL semantics, and bearing on [Partee & al 90 chap 14] and [van Benthem 86 Part I] another one which is truly empirical and which deals with monotonicity of different kind of determiners and their possible combination with negation.

Grammaire de Montague et ambiguïté

Résumé

Les questions traitées (§1) portent sur les deux points suivants.

- Est-ce que, dans la syntaxe de la Grammaire de Montague, on peut avoir des expressions graphiques, élémentaires ou non, associées à plus d'une catégorie ?
- Si la réponse à la question précédente est oui, quelles sont les conséquences que cela entraîne pour les autres pièces du système de la Grammaire de Montague, en particulier pour la traduction en LI et/ou l'interprétation de la syntaxe dans un univers de dénotation ?

Elles sont abordées par rapport à UG (§ 2) et par rapport à PTQ (§ 3). On suppose que R est utilisée comme une fonction qui consiste à effacer les parenthèses indexées introduites par les opérations d'un langage non ambigu.

L'ambiguïté d'assignation catégorielle dans UG est formellement définie. Elle est formellement possible (§ 2.1), de telle manière que l'on peut avoir, selon les définitions de UG, un langage non ambigu avec des expressions, élémentaires ou non, appartenant à des ensembles indexés différemment par des indices catégoriels appartenant à Δ . Des manières possibles d'obtenir des assignations catégorielles ambiguës pour des expressions non élémentaires sont explicitées. On explicite aussi et on donne des exemples de la relation entre l'ensemble noté EBF (Expressions Bien Formées) et l'ensemble noté $E \sim A$ (Expressions non Ambiguës, notant en abrégé l'ensemble $\cup_{\delta \in \Delta} C\delta$ qui ne reçoit pas en UG de nom particulier), le premier étant obtenu par effacement par R des parenthèses indexées dans les expressions des seconds. Les répercussions systémiques de la non-unicité d'assignation catégorielle sont analysées par rapport aux définitions de UG d'ambiguïté syntaxique (§ 2.2.1), d'ambiguïté sémantique (§ 2.2.2) et d'ambiguïté de traduction (§ 2.2.3), cette dernière notion n'étant pas explicitée dans UG mais pouvant s'extrapoler de la précédente. Les conclusions sont :

- L'unicité d'assignation catégorielle n'est pas une exigence formelle sur les langages non ambigus, mais la non-unicité d'assignation catégorielle d'une expression ne permet pas d'exprimer une ambiguïté syntaxique de cette expression.
- Comme une expression ne peut être sémantiquement ambiguë que si elle est syntaxiquement ambiguë, une expression, du fait de son appartenance à des ensembles indexés différemment, ne peut pas être considérée sémantiquement ambiguë ; la non-unicité d'assignation catégorielle d'une expression ne permet donc pas d'exprimer une ambiguïté sémantique de cette expression.
- L'existence d'expressions, élémentaires ou non, appartenant à des ensembles indexés différemment engendre pour le système un non-sens formel.
- *Mutatis mutandis*, on doit aboutir aux mêmes conclusions pour l'ambiguïté de la traduction.

Les questions traitées doivent être analysées différemment dans le cadre de PTQ ; on fait le rappel des différences par rapport à UG (§ 3). On constate que, ni dans PTQ ni dans aucun des travaux consultés l'ayant présenté ou commenté, il n'y a d'assignations catégorielles ambiguës.

Sans doute, la possibilité formelle du non-respect des assignations catégorielles uniques subsiste en PTQ (§ 3.1). Les répercussions systémiques sont difficiles à établir étant donné l'indéfinition formelle des arbres de dérivation utilisés dans PTQ, et l'opacité de la relation entre expressions d'un langage non ambigu à la UG et arbres de dérivation à la PTQ. Sur un point crucial – la possibilité d'avoir des nœuds indexés dans l'arbre –, l'interprétation de [Chambreuil 89], qui les accepte, est différente de leur non-utilisation dans PTQ.

Les conclusions auxquelles on arrive sur les répercussions systémiques dans PTQ de la non-unicité d'assignation catégorielle sont les suivantes :

- Pour les expressions élémentaires, les conclusions sont analogues à celles de UG : on aboutit à un non-sens formel. Si pour pallier ce non-sens on introduit des stipulations ad hoc, on obtient des résultats descriptivement inadéquats : chaque fois que par assignation ambiguë il y a des solutions concurrentes, si l'on choisit arbitrairement l'une, le résultat sera erroné pour les expressions qui en réclament une autre.

PTQ ne définit pas formellement la différence entre arbres de dérivation « authentiquement » différents – c'est-à-dire conduisant à des interprétations sémantiquement différentes –, et arbres de dérivation qui sont différents mais qui ne le sont pas « authentiquement ». Cependant on peut accepter que cette distinction soit opérationnelle. On peut aussi admettre, avec PTQ, qu'un arbre de dérivation admet une seule traduction. La question est donc de savoir si une expression appartenant à des ensembles indexés différemment est associée de ce fait à des arbres dérivationnels « authentiquement » différents. On aboutit à deux réponses :

- Il est formellement impossible d'associer cette expression à des arbres dérivationnels « authentiquement » différents, si ces arbres sont obtenus à la façon Montague.
- Il est formellement possible d'associer cette expression à des arbres dérivationnels « authentiquement » différents, si ces arbres sont obtenus à la façon Chambreuil et à condition de respecter l'unicité d'assignation des expressions élémentaires.

Montague Grammar and ambiguity

Abstract

The two approached issues (§ 1) are :

- Is it possible in the syntax of Montague Grammar to have an expression, basic or not, associated to more than one category?
- If the answer to the previous question is 'yes', which are the consequences for the other elements in the overall system of Montague Grammar, and in particular with respect to the translation into IL and/or to the interpretation of syntax in the denotational universe?

These questions are approached in the UG perspective (§ 2) and in the PTQ one (§ 3). R is used as a function which is intended to delete the indexed parentheses introduced by the operations of a disambiguated language.

Given, in UG terms, a formal definition of ambiguous categorial assignement, the conclusion is that it is formally possible (§ 2.1) both to assign some expression, basic or not, to sets differently indexed by indices from Δ , and to satisfy the UG definition of a disambiguated language. Ways to obtain ambiguous categorial assignement of non-basic expressions are explicited. The relation thus arising between the set noted EBF (in French *Expressions Bien Formées*, in UG *meaningful expressions*) and the set noted $E \sim A$ (in French *Expressions non-Ambiguës*) is explicited. In UG this last set does not receive a particular name and is noted $\cup_{\delta \in \Delta} C\delta$; the former set is obtained by deletion by R of the indexed parentheses in the expressions of the latter one. The systematic repercussions of the non-uniqueness of categorial assignment are examined in relation to UG definitions of syntactic ambiguity (§ 2.2.1), of semantic ambiguity (§ 2.2.2) and of translation ambiguity (§ 2.2.3). This last notion is not explicited in UG, but it can be extrapolated from the previous one. The conclusions are :

- Uniqueness of categorial assignement is not a formal requirement on disambiguated languages, but the non-uniqueness of categorial assignment of some expression does not allow to express a syntactic ambiguity of this expression.
- Because an expression can be semantically ambiguous only if it is syntactically ambiguous, an expression assigned to more than one set differently indexed cannot be considered semantically

ambiguous; thus, the non-uniqueness of categorial assignement of some expression does not allow to express a semantic ambiguity of this expression.

- The existence of the same expression, basic or not, in sets differently indexed, results in a formal nonsense of the system.
- *Mutatis mutandis* the same conclusions are obtained for ambiguity of translation.

The targeted questions are differently analyzed in a PTQ perspective. Differences with respect to UG are recalled (§ 3) and the following fact is established : there are no ambiguous categorial assignements neither in PTQ nor in the consulted studies presenting or commenting it.

Without any doubt, the formal possibility of non-uniqueness of categorial assignements remains in PTQ (§ 3.1). The systematic repercussions of the non-uniqueness of categorial assignement are difficult to establish given the PTQ loose formal definition of analysis trees and the opacity of the relation between expressions of a disambiguated language conceived in a UG manner and PTQ analysis trees. On a crucial point, the possibility of having indexed nodes in the analysis tree, there is an important divergence between [Chambreuil 89], who accepts them, and PTQ, that does not use them (§ 3.2).

The conclusions on PTQ systematic repercussions of non-uniqueness of categorial assignements are the following (§ 3.2) :

- With respect to basic expressions, the conclusions are analogous to the ones related to UG : ambiguous categorial assignement of expressions results in a formal nonsense of the system. If ad-hoc stipulations are introduced in order to save this formal nonsense, the results that are obtained are descriptively inadequate. Every time that, because of ambiguous categorial assignement, there are two or more possible solutions, if one of them is arbitrarily chosen, inadequate results will be obtained for expressions requiring another one.

PTQ does not formally define the difference between analysis trees “genuinely” different – i.e. which must be associated to different semantic interpretations – and analysis trees which are different but not “genuinely” different. However, it is possible to safely assume that this distinction works. It is also possible to admit, as in PTQ, that an analysis tree accepts only one translation. The question is to know whether an expression belonging to differently indexed sets will be associated to “genuinely” different analysis trees. Two different answers are obtained :

- It is formally impossible to associate this expression to “genuinely” different analysis trees, if the trees are obtained in a Montague way.
- It is formally possible to associate this expression to “genuinely” different analysis trees, if the trees are obtained in a Chambreuil way, inasmuch the uniqueness of categorial assignement of basic expressions is respected.

Empiricité en linguistique et grammaire de Montague :

la sémantique en 5P et la compositionnalité

Sommaire

1 Objectifs	5
1.1 Difficulté de la tâche	
2 Apports de la Grammaire de Montague	12
3 Les interrogations de ce travail en fonction des objectifs	15
3.1 L'interrogation sur l'empiricité	
3.2 L'interrogation sur la compositionnalité	
3.2.1 La compositionnalité dans UG	
4 Syntaxe de la langue naturelle et formules de la logique intensionnelle dans PTQ	33
4.1 Le langage formel (pSLF)	
4.1.1 Notions utilisées	
4.1.2 Spécification des couples <règle-pSLF, opération-pSLF> de PTQ	
4.1.3 Le pouvoir expressif des couples-pSLF	
4.1.4 Le Vocabulaire de pSLF	
4.1.5 Permissivités et restrictions de pSLF	
4.2 La traduction du langage-pSLF en langage-LI	
4.2.1 Les restrictions sur la traduction	
4.2.2 Les permissivités de la traduction	
4.3 Analyse d'exemples	
5 Bilan sur le traitement de LN dans PTQ	53
5.1 Notation	
5.2 Reconsidération des exemples 1 à 3 de la Section 4.3	
5.2.1 Syntagmes prépositionnels comme modificateurs des noms	
5.2.2 Lemmes verbaux avec plus d'une sous-catégorisation	
5.2.3 Emplacement des adverbes	
5.3 La généralisation des solutions	
5.3.1 Panoplie des principales permissivités utilisées et de leur rôle	
5.3.2 L'obtention des solutions	
5.3.2.1 L'exploitation de la multiplicité des slashes	
5.3.2.2 Ambiguïté catégorielle des expressions terminales	
5.3.2.3 Relations d'ordre	
6 Éléments pour une réévaluation de la Grammaire de Montague et de ses interprétations	69
7 Empiricité en linguistique et homomorphisme syntaxe-univers de dénotation.	77
7.1 Du sel de cuisine et du chlorure de potasse ou de l'utilisation empirique des morphismes algébriques	
7.2 L'homomorphisme syntaxe-univers de dénotation et sa possible utilisation empirique ou la naturalité de la langue naturelle	
8 5P et compositionnalité	88
Notes complémentaires	103
Références	109

1 Objectifs

L'objectif essentiel et général de ce document est de justifier en amont les choix du Paradigme 5P (désormais 5P, cf. un résumé succinct au § 8) sur la description de la sémantique. En particulier, il s'agit :

- d'aider à prendre position sur les contributions de Montague et sur ce qu'on connaît en général comme *Grammaire(s) de Montague* ;
- de mieux spécifier des fonctionnalités qu'on veut assigner aux Fonctions Sémantiques en vue de calculer les Représentations sémantiques à partir du graphe déterminé par l'ensemble de couples de fléchage associé à chaque phrase et des entrées du lexique avec leur système de subsumption et d'héritage entre catégories ;
- d'aider à caractériser 5P par rapport à l'existant et ceci en fonction du choix méthodologique de base de 5P qui consiste à essayer de pratiquer effectivement la linguistique comme une science empirique.

Pour atteindre ces objectifs, il a été nécessaire de passer par une révision approfondie du rapport syntaxe-sémantique chez Montague. De ce fait, le document se concentre ainsi (cf. Section 4) sur :

- l'objet dit *langage formel* dans PTQ utilisé pour décrire les suites de la langue naturelle (désormais *langue naturelle* se note *LN* ; pour l'interprétation des sigles *EFL*, *UG*, *PTQ*, cf. ci-dessous Références)
- la traduction – utilisant la terminologie de Montague – des suites de *LN* en formules de la logique intensionnelle (désormais *logique intensionnelle* se note *LI*).

Ce document a dû ainsi aborder les notions de compositionnalité de la sémantique, de parallèle syntaxe-sémantique, etc. et ceci dans le cadre des affirmations attribuées à Montague sur la non-spécificité de la langue naturelle par rapport aux langages artificiels (cf. Section 3). Il ne traite en rien les opérations de déductibilité sur les formules *LI* dans la *LI* proposée par Montague ni l'univers de dénotation, ni a fortiori l'interprétation des formules *LI* dans l'univers de dénotation. Sa compréhension ne nécessite en rien de la connaissance du formalisme de 5P.

1.1 Difficulté de la tâche

Il n'est pas sans difficulté d'essayer d'atteindre les objectifs visés étant donné une absence de clarté :

- sur des notions méthodologiques et linguistiques permettant de poser les problèmes ;

- sur la délimitation du domaine abordé et sur les relations internes aux objets de ce domaine.

Le premier type de difficulté peut être illustré par les termes *hypothèse*, *formel* et *langage*. Le deuxième par les objets dont on parle avec les notions de *Grammaire(s) de Montague* et *principe frégeen* et/ou *compositionnalité*.

Il serait facile de multiplier les citations pour montrer à quel point des expressions comme *théorie*, *hypothèse* et autres relevant de la méthodologie des sciences sont extraordinairement démonétisées en linguistique, et cela en linguistique standard ou classique et en linguistique formelle et/ou informatique. Dans ce document, *théorie* et *hypothèse* sont des expressions qui dénotent des formulations relevant d'une science du réel, ou de l'empirie, dans la terminologie de [Granger 92], des formulations qui doivent être susceptibles d'être testées – ou qui doivent être falsifiables dans une terminologie popérienne – et qui, après test, peuvent être vérifiées ou non vérifiées, en ne discutant pas l'idée maintenant bien ancrée (cf. [Sokal & Bricmont 97 chap. 3]) que d'une hypothèse ou d'une théorie on ne peut pas dire qu'elle est vraie, seulement que jusqu'à un moment déterminé, on n'a pas réussi à montrer qu'elle est fausse. Pour qu'une formulation acquière le statut d'hypothèse, et qu'un ensemble d'hypothèses acquière le statut de théorie, il faut que théorie et hypothèses soient associées à une manière précise de faire des observations et à une manière précise de comparer les déductions que l'on peut tirer des hypothèses avec les observations, représentées – elles aussi – d'une manière précise. N'importe quelle idée, suggestion, principe d'organisation, guide intellectuel, etc. n'est pas une hypothèse – dans l'utilisation que nous faisons ici de ce terme et de ce concept. Et il est inutile de se demander, sur un plan général, où passe la limite exacte entre les idées et/ou les suggestions d'une part, et les hypothèses de l'autre, étant donné que l'on abordera des situations où strictement il n'y a quasiment rien qui ait été explicité et qui puisse servir d'observation susceptible de tester une hypothèse.

Des observations analogues peuvent être faites sur *formel*. Nous réservons *formel* pour caractériser des formulations exprimées dans un langage avec une syntaxe et sémantique explicites, formulations sur lesquelles on peut faire des calculs et dont on connaît les propriétés mathématiques. Un énoncé quelconque ne devient donc pas *formel* du seul fait de le caractériser ainsi ou de l'exprimer moyennant flèches ou symboles permettant une notation compacte.

Pour en finir avec le premier type de difficultés, revenons à *langage* et aux termes de *grammaire*, d'*expression* et/ou *expressions élémentaires* et/ou d'*expressions significatives*, toutes expressions qui d'une manière ou d'une autre se rapportent, en se distinguant, au *langage*.

Le terme anglais *language*, dans des utilisations non techniques, est utilisé soit pour désigner un ensemble d'expressions soit pour désigner une organisation supposée sous-

jacente à un ensemble d'expressions. Dans les utilisations plus techniques, le terme *grammar* est en général employé pour désigner cette supposée organisation sous-jacente. Le problème est que Montague, dans des écrits techniques, utilise *language* d'une manière complexe (cf. ci-dessous) qui comprend l'organisation sous-jacente elle-même. Cela donne lieu à de mauvaises interprétations, dont certaines signalées dans [Dowty et al. 81 p. 255], où on remarque que, à la place de *disambiguated language*, expression utilisée par Montague, l'expression *unambiguous grammar* « would be a more natural term for ».

Dans le sillage de cette confusion terminologique autour de *langage*, nous en trouvons d'autres. Dans [Dowty & al. 81] on trouve – malheureusement pour le lecteur dans le dernier chapitre (p. 252) – un autre *caveat* terminologique important, et cette fois sur *grammaire universelle*.

A word on terminology is in order to begin with, because Montague used the term "universal grammar" in a different sense than the one it has in linguistics. In the latter field, universal grammar is an account of those features that are common to all natural languages –in particular of the essential characteristics of natural languages, which distinguish them from other logically possible systems of communications such as artificial languages used in mathematics or computer programming, animal "languages" like that of bees, etc. In contrast to this conception, Montague used the term to refer to a mathematical framework sufficiently general to subsume a description of any system that might conceivably be termed a language, whether it is a possible natural language or not. Thus it ought to be possible in Montague's universal grammar to give a description of the Fortran programming language or of bee "language" just as readily as we were to describe a fragment of English, perhaps even more so in view of the relative simplicity of these languages.

Enfin, et pour en finir avec cette constellation de confusions terminologiques, reprenons *expression*, *expression élémentaire*, *expression significative*. Supposons un ensemble de symboles graphiques, p. ex. les lettres de notre alphabet. Supposons des suites de lettres, que nous nommons *expressions graphiques* (et faisons cela pour ne pas retomber dans l'ambiguïté de *mot*). Ainsi, *lettre*, *aimable*, *toit* sont des expressions graphiques. Supposons une opération de concaténation sur des expressions graphiques, opération qui intercale le symbole graphique « - » entre deux expressions graphiques. L'expression *lettre-aimable* serait le résultat d'une telle opération sur les expressions graphiques *lettre* et *aimable*. Nommons *expressions élémentaires* les expressions graphiques qui ne résultent d'aucune opération, de concaténation ou autre. Supposons toujours un ensemble d'expressions graphiques élémentaires et un ensemble d'opérations. Étant donné ces deux types d'éléments, il y aura des expressions qui peuvent être formées par ces opérations – comme p. ex. *lettre-aimable* – et d'autres qui ne pourront pas l'être – comme vraisemblablement, p. ex. *lettre-toit*. Celles qui peuvent l'être sont nommées parfois *expressions significatives*, ou, dans une autre terminologie, mais voulant dire la même chose, *expressions bien formées*.

Les opérations sur les expressions graphiques ne se définissent pas en général en utilisant des énumérations d'expressions graphiques. Mais on classe plutôt les expressions graphiques, et on les associe à ce qu'on nomme souvent une *catégorie*, terme qui introduit

une autre ambiguïté terminologique, sur laquelle on revient ci-dessous. Donnons pour l'instant à *catégorie* l'interprétation d'étiquette. Acceptons que l'on peut associer aux expressions – élémentaires ou non – des catégories, au sens que l'on vient de fixer. Le problème en est que par *expression* on ne sait pas souvent si l'on a un objet graphique associé à une catégorie, ou un objet graphique, pour ainsi dire, nu. Est-ce que les expressions graphiques élémentaires correspondant à *le* et *chien* sont respectivement *le* et *chien* ou elles sont *le*_{Art} et *chien*_{Nom}?

Lorsqu'on parle de langages formels, on entend toujours par *expression* expression graphique nue. Si l'on dit *une expression a ceci ou cela* le ceci et cela en question sont exprimés dans le graphisme de l'expression. Tel n'est pas toujours le cas lorsqu'on parle des *expressions* de la langue naturelle, en omettant de dire si ce dont il s'agit sont ou non des expressions graphiques nues, ou des expressions graphiques avec un plus sous forme de catégorie. Et cette distinction entre les deux *expressions* va être cruciale pour comprendre quel type de relation existe entre expressions de la langue naturelle et l'univers de dénotation.

Même au risque d'alourdir le texte, nous utiliserons systématiquement :

- *suite* (éventuellement d'un seul élément) ou *expression graphique* pour dénoter une expression graphique nue, sans que le graphisme de l'expression soit associé à quelque chose ;
- *langage-LN*, pour dénoter un ensemble de suites de LN ;
- *SxLy*, soit *Spécification (d'un) Langage*, pour dénoter l'organisation sous-jacente à un langage; la variable notée par *x* entre *S* et *L* sera instanciée par *F* (formelle, par rapport à *S*) ou *~F* (non formelle); la variable notée par *y* à droite de *SL* sera instanciée par *A* et *~A* (par rapport à *Langage*) pour noter *SxLA* (Spécification d'un Langage Ambigu) et *SxL~A* (Spécification d'un Langage non Ambigu) ;
- *expression*, pour dénoter soit une suite, soit une suite associée à une étiquette ;
- *langage*, pour dénoter un ensemble d'expressions ;
- *langage-SxLy*, pour dénoter l'ensemble d'expressions spécifiées par *SxLy* ;
- *form-LI*, pour dénoter une expression graphique de LI ;
- *langage-LI*, pour dénoter un ensemble de form-LI.

On conserve ainsi *expression* (et *expression élémentaire*, *significative*, etc., et *langage* sans suffixe) pour une utilisation ambiguë, ce qui exigera de se poser le problème de savoir si l'on pense *suite* (c'est-à-dire graphisme nu), ou si, en fait, on veut dire *suite avec étiquette*, et dans ce cas, laquelle. Pour compliquer encore la situation il faut avoir à l'esprit que les expressions *langage* et *langage non ambigu* chez Montague (cf. Section 3.2.1) doivent se comprendre par une notion résultant d'un panachage des deux notions ici distinguées : en partie la notion de *langage*, en partie celle de *Spécification d'un Langage*, telles que nous venons de les caractériser. Pour cette raison on va noter toujours « *langage* » et « *langage non ambigu* » – entre guillemets – pour faire référence aux notions montaguiniennes.

Revenons à *catégorie*. Supposons un ensemble d'expressions graphiques, élémentaires ou non, qui doivent, toutes, être associées à la même *catégorie*, au sens d'étiquette. Le problème est qu'on utilise le terme *catégorie* pour dénoter et l'étiquette elle-même et l'ensemble d'expressions qui doivent être associées à une même étiquette. Ceci par Montague lui-même, qui a pourtant bien pris la peine, lui, d'avertir qu'il parlait de choses différentes. Ainsi dans la première des citations suivantes, de UG, avec *catégories* il parle d'un ensemble d'expressions, alors que, dans la deuxième, de PTQ, il parle des étiquettes, en introduisant l'idée de les définir récursivement à partir des symboles primitifs *e* et *t* :

The basic aim [...] of syntax is to characterize the various syntactical categories, especially the set of declarative sentences (UG p. 223 note 2).

It should be pointed out that our categories are not sets of expressions but will instead serve as *indices* of such sets (PTQ, p. 249).

En plus de cette absence de clarté que nous venons de voir dans les notions méthodologiques et linguistiques permettant de poser les problèmes, nous avons identifié, au tout début de cette section, une autre absence de clarté, celle-ci portant sur la délimitation du domaine abordé et sur les relations internes aux objets de ce domaine.

La première difficulté vient de la délimitation de ce qu'on entend par *Grammaire(s) de Montague*. En effet, cette expression est remarquablement floue. Comme [Partee 97, p. 7] le signale, elle « may be token in a narrower or a broader sense, since continuing research has led to a variety of work that can be considered to invoke either 'developments of' or 'departures from' Montague's original theory and practice [...] the boundaries are vague [...] The term has never been restricted to Montague's work alone and it should not be ».

Ce document ne satisfait pas à l'injonction de Partee. Il utilise l'expression *Grammaire de Montague* d'une manière hyperétroite et restrictive : avec cette expression on ne dénote que le travail de Montague dans UG et PTQ.

Il se pose ensuite et de manière cruciale la question de la délimitation de la notion du principe de compositionnalité et/ou de principe frégeen (auquel il y a des allusions dans UG, allusions qui peuvent ou doivent s'interpréter d'une manière assez restreinte, cf. Section 3.2), qui sont typiquement des notions fonctionnant comme une auberge espagnole d'autrefois : on y trouve à peu près ce qu'on y apporte.

On se réfère ici pour justifier l'assertion précédente aux travaux de Janssen et en particulier à [Janssen 97], dont on extrait les citations suivantes; la première concerne la paternité douteuse de Frege du principe qu'on lui attribue, la deuxième porte sur l'indéfinition du principe de *compositionnalité* si on laisse ouverte la caractérisation des éléments essentiels de la définition (cette indéfinition sera reprise dans la Section 3.2).

The conclusion is that Frege rejected the principle of compositionality in the period in which he wrote *Grundlagen des Mathematik* but may have accepted the principle later on in his life. It seems that nowhere in his published works does he mention compositionality as a principle. It is, therefore, inaccurate to speak of 'Frege's principle'. Compositionality is not Frege's, but it might be called 'Fregean' because it is the spirit of his later writings (p. 421).

[...] because the formulation of the principle is sufficiently vague, that anyone can put his own interpretation on the principle (p. 419).

Il se pose enfin la question des rapports complexes entre définitions de EFL et de UG et présentations de PTQ. On sait que Montague est passé d'un système en trois éléments dans EFL, synthétisé dans (i), à un système en cinq, synthétisé en (ii), utilisé déjà dans UG.

- i
 - SFL, censé spécifier le langage-LN ;
 - fonction interprétative, associant le « langage »-SFL à l'univers d'interprétation ;
 - univers d'interprétation.
- ii
 - SFL, toujours censé spécifier le langage-LN ;
 - langage-LI ;
 - univers d'interprétation ;
 - traduction du « langage »-SFL en langage-LI ;
 - fonction interprétative entre le langage-LI et l'univers d'interprétation.

Cette présentation, claire et synthétique, est empruntée à [Chambreuil & Pariente 90 p. 66 et p.163], mais en y introduisant des modifications. Celles-ci tiennent, d'une part, à l'utilisation de la terminologie présentée ci-dessus (p. ex. dans le texte original, on lit *langage formel* là où on a utilisé *SFL*) ; d'autre part, c'est nous qui introduisons l'idée que *SFL* est censé spécifier le langage-LN, idée qui sera reprise et justifiée dans la Section 4, car, en fait, d'après nous, la spécification de (ii) dans PTQ relève plutôt de *S~FL* que de *SFL*.

Ce qu'il nous importe de remarquer ici c'est une complexité des rapports entre (i) et (ii), complexité qui ne se dégage pas forcément de manière immédiate de la lecture des deux points précédents ainsi présentés. Comme illustration de cette complexité, on remarquera que dans UG (cf. Section 3.2.1) est définie la contrainte d'homomorphisme entre algèbres – clef d'une interprétation possible de la compositionnalité, cf. Section 3.2 – et que l'algèbre qui aura un rapport avec la LN est placée dans un *SL* qui est un *SL~A*, alors que le *SL* de PTQ est un *SLA*. En fait, UG définit *SLA* comme un couple $\langle \langle SL~A \rangle, R \rangle$, alors que la relation *R* (celle qui va permettre de relier *SL~A* à *SLA*) n'est pas contrainte et l'homomorphisme est exigé par rapport aux deux premiers éléments du quintuplet qui définit *SL~A* et l'algèbre qui définit l'univers de dénotation. En PTQ, le fragment observé de l'anglais est décrit par un *S~FLA*, et c'est le « langage »-*S~FLA* qui est traduit en

langage-LI¹. Et cette traduction s'effectue par des éléments qui caractérisent la base de traduction définie en UG, et non pas, *stricto sensu*, par cette base. Cela est clairement indiqué dans [Chambreuil 89 p. 78] :

Nous présenterons ci-dessous [...] les éléments d'un système de traduction du fragment de l'anglais PTQ sur le langage de la logique intensionnelle. Éléments d'un système de traduction et non pas, plus formellement, base de traduction, simplement parce que le fragment façon PTQ n'est pas caractérisé en tant que langage [dans notre terminologie, on dirait *n'est pas caractérisé en tant que S~FL~A*]. Pour caractériser une base de traduction il faudrait travailler avec LQ [LQ est présenté dans [Chambreuil 89 p. 56-56] ; dans notre terminologie LQ correspond à une spécification du fragment de l'anglais de PTQ via <<S~FL~A >, R>], et nous avons choisi de suivre ce qui est devenu tradition.

Pour compléter le cadre des difficultés, il est impossible de ne pas faire état de la difficulté première, c'est-à-dire l'accès direct aux textes originaux de Montague, qui sont, comme l'auteur le reconnaît lui-même, plutôt hermétiques :

For the sake of brevity I shall content myself with the mere statement of definitions [...] avoiding almost all discussion and intuitive amplification. The resulting exposition will, I realize, be cryptic and unsatisfactory [...] (UG p. 223).

Il est probable que, dans l'histoire des sciences du langage, les 48 pages qui occupent UG et PTQ dans [Montague 74] resteront celles qui, proportionnellement, auront déclenché le plus grand nombre de pages d'introductions, de commentaires, de développements. Heureusement il existe des ouvrages excellents qui introduisent aux textes de Montague. Ils ont été largement utilisés dans ce travail. En fait il serait légitime de dire que l'on présente ici une vision de Montague, centrée sur PTQ avec les incursions nécessaires dans UG pour répondre à nos interrogations, via [Halvorsen & Ladusaw 79], via [Dowty & al. 81], et via les travaux de Michel Chambreuil, qu'il a écrits seul ou en collaboration, à savoir :

Chambreuil 89
Chambreuil & Pariente 90
Chambreuil 91
Chambreuil 98
Chambreuil & al. 98

[Halvorsen & Ladusaw 79], [Dowty & al. 81] et les travaux de Chambreuil sont largement et principalement utilisés dans les Sections 3 et 4. Dans les Sections 6 à 8 on a élargi les références afin de mieux positionner le travail par rapport à d'autres contributions qui se sont penchées sur la relation syntaxe-sémantique dans une optique montaguienne.

Les travaux d'introduction utilisés ont été largement utiles, mais la panoplie des difficultés ne serait complète sans mentionner que, au-delà des divergences habituelles en notation, emphases différentes sur tel ou tel aspect de l'œuvre montaguienne, on trouve aussi, dans

¹ Cf. dans la Section 4.1 la justification de l'utilisation de ~F.

ces travaux d'introduction, des divergences importantes sur des points particuliers. Ainsi par exemple, la question de l'association, possible ou non, d'une même expression graphique élémentaire à deux catégories différentes dans un SFL~A. D'après [Chambreuil & al. 98] c'est possible, non seulement pour les expressions graphiques élémentaires mais aussi pour les expressions graphiques qui ne le sont pas (cf. la première des citations suivantes) sans aucune condition, alors que [Dowty & al. 81] ne fait état de la possibilité d'appartenance d'une même expression à plus d'une catégorie que si l'expression est élémentaire et qu'elle satisfait une condition – le fait que l'expression ait « the same meaning regardless of category membership » (cf. la deuxième citation). Et les deux travaux parlent, bien évidemment, d'un même objet, la SFL~A telle qu'elle est présentée dans UG :

[...] rien n'interdit [...] qu'une même expression significative (en particulier les expressions élémentaires) appartienne à plusieurs catégories syntaxiques [Chambreuil & al. 98 p. 43].

A set of basic expressions apportioned into syntactic categories [...] he [c'est-à-dire, Montague] allows a basic expression to belong to more than one category but will require that it have the same meaning regardless of category membership [Dowty & al. 81, p. 255].

Et cette question n'est pas sans importance, car derrière elle il y a toute la problématique de l'ambiguïté d'assignation catégorielle aux expressions de la LN. Elle n'est pas d'un abord facile, étant donné les interprétations possibles de *expression* et de *catégorie*.

Pour essayer de faire face au mieux à cet ensemble hétérogène des difficultés, le choix a été fait d'essayer d'explicitier au maximum, au risque d'une possible redondance, de multiplier les citations explicites, en pointant parfois de manière analytique, dans les textes cités, la justification de telle décision ou de telle interprétation, de présenter avec un souci d'explicitation la plupart des exemples étudiés dans la Section 4.3 et 5. Après une première présentation synthétique des apports de Montague dans la Section 2 – suivante – on présentera dans la Section 3, de manière plus détaillée, les interrogations centrales de ce travail et le schéma dans lequel elles ont été traitées.

2 Apports de la Grammaire de Montague

On peut organiser les apports de la Grammaire de Montague autour des trois pôles suivants :

- i Idées générales et profondes qui se situent à la lisière de la logique, de la linguistique et de la philosophie du langage.
- ii Spécification syntaxique des LN et traduction des suites-LN en form-LI.
- iii Définition d'une logique intensionnelle (désormais LI) avec son univers de dénotation et ses règles de calculabilité.

La LI de (iii) a notamment permis de traiter formellement – i.e. en se donnant des univers de dénotation, et des règles de déductibilité – des observations qui ont cours en philosophie du langage sur les « lectures spécifiques », « lectures référentielles », « lectures intensionnelles » ou « lectures extensionnelles ». Il a permis de traiter ce que Quine a baptisé la « quantification in » pour traiter le problème classique de *Jean ne sait pas que la capitale du Nicaragua est Managua*, qui exige des valeurs de vérité contingentes en fonction de l'état du monde, tandis que la phrase *Jean ne sait pas que Managua est Managua* semble irrémédiablement fausse pour tout état du monde, alors que *la capitale de Nicaragua* et *Managua* ont la même dénotation. Dans ce document, on n'examinera en rien cet apport fondamental de la Grammaire de Montague.

En revanche ce document se propose de revisiter les idées montaguiniennes du pôle (i). Elles sont synthétisées dans cette Section 2. Dans la Section 3 on explicitera les deux angles principaux à partir desquels elles seront examinées, à savoir leur empiricité et leur exigence de compositionnalité. Ces idées du pôle (i) seront reprises pour en donner une évaluation dans les Sections 3 – en ce qui concerne UG –, dans les Sections 4 et 5 pour ce qui concerne PTQ. Dans la Section 6 on donnera une vision d'ensemble, en procurant de synthétiser des éléments qui peuvent contribuer à une ré-évaluation de la Grammaire de Montague. La Section 7 suggère une possible utilisation empirique des idées montaguiniennes et la Section 8 fixe la position de 5P par rapport à la compositionnalité. Pour aboutir aux conclusions des Sections 6 à 8, il a été nécessaire d'analyser dans le détail les points concernant le pôle (ii), ce qui est fait dans les Sections 3.2.1, 4 et 5.

En discours rapporté, on trouve la meilleure synthèse des idées du pôle (i) dans [Chambreuil & Pariente 90]. Elles sont ici transcrites dans la suite par des citations textuelles :

- Idée 1 Il n'y a pas de différences théoriques fondamentales entre les langages naturels et les langages artificiels des logiciens; la syntaxe et la sémantique de ces différents langages relèvent d'une même théorie mathématique (p. 19).
- Idée 2 L'intérêt principal de la syntaxe est de conduire à la sémantique (p. 64).
- Idée 3 L'objectif principal de la sémantique est de caractériser les notions d'énoncé vrai et de conséquence (p. 64).

Les Idées labellisées 1 à 3 ci-dessus sont reprises dans [Chambreuil 98, p. 44-45] sous des formulations légèrement différentes, dans la section qui porte comme titre « Les hypothèses générales : l'objet de la sémantique des langues naturelles ». Elles y sont développées ainsi :

... les logiques, dans leurs versions formelles du vingtième siècle, sont traditionnellement définies à partir de trois notions générales principales : celle de système formel, celle d'univers d'interprétation (ou de modèle), celle d'interprétation qui fait passer du système formel à l'univers d'interprétation. Ce sont ces trois notions qui sont généralisées, définissant les éléments de la théorie mathématique commune aux logiques et aux langues naturelles. Un système formel et un univers d'interprétation sont des cas particuliers d'une structure mathématique plus générale. Chacun des deux est une algèbre, c'est-à-dire une structure caractérisée par un ensemble et une

famille d'opérations définies sur cet ensemble. Une interprétation et une représentation peuvent alors être définies à leur tour comme des fonctions entre ces deux structures, fonctions auxquelles peuvent être associées différentes propriétés, par exemple celle de *préserver* les structures mises en correspondance. Dans ce sens, des opérations de l'univers d'interprétation sont associées à celles de l'univers formel. Alors, si sur le système formel une expression est obtenue comme résultat d'une opération appliquée à d'autres expressions, son interprétation sur l'univers d'interprétation est obtenue comme résultat de l'opération associée à l'opération du système formel, appliquée aux interprétations des expressions arguments de cette opération. [...] Pour une langue naturelle, cela se traduira, en particulier, par les éléments suivants. Par exemple, une entité de l'univers d'interprétation sera associée à « Jacques », une autre entité sera associée à « parler ». Une opération (syntaxique) permettra de construire « Jacques parle ». A cette opération correspondra une opération sur les entrées associées respectivement à « Jacques » et à « parler » qui (hors problème de temporalité) appliquée à ces entités aura pour résultat l'entité de l'univers d'interprétation associée à « Jacques parle ». Une logique et une langue naturelle sont alors des cas particuliers d'une architecture constituée par un langage (formel) et un univers d'interprétation, chacun d'eux ayant une structure d'algèbre, et par une fonction d'interprétation préservant les deux structures, au sens évoqué ci-dessus, et définissant la mise en correspondance entre les expressions du langage formel et les entités de l'univers d'interprétation (Chambreuil 98, p. 45-46).

On a ici une présentation formellement claire d'une des versions de la compositionnalité sous la forme d'un homomorphisme entre structures algébriques de la langue naturelle et l'univers de dénotation. La clef de cette vision du problème réside bien dans la préservation des structures algébriques en langue naturelle et dans l'univers d'interprétation par la fonction d'interprétation, ce qui est une formulation claire du « parallélisme syntaxe-sémantique ». Le problème central n'est donc pas celui de se donner un système explicite de construction d'une représentation sémantique quelconque à partir d'une suite d'expressions terminales, mais bien celui de décrire la langue naturelle par le biais de structures algébriques :

Le problème n'est donc plus ici de *traduire* les expressions d'une langue naturelle dans des expressions du langage d'une logique en ne retenant de la langue naturelle que des régularités correspondant à celles qui ont été retenues pour la définition de la logique prise comme espace de traduction, et on peut penser par là, par exemple, aux traductions en logique des propositions ou en logique des prédicats. Le problème est de caractériser une langue naturelle en elle-même, selon des mécanismes mis en œuvre pour définir une logique. La granularité d'analyse des régularités doit correspondre à la granularité des constituants de la langue naturelle (Chambreuil 98, p. 46).

Et les structures algébriques de la langue naturelle, captées moyennant SFL (langage formel, dans les citations) doivent être homomorphiques avec celles de l'univers d'interprétation :

[...] l'association d'une entité de l'univers d'interprétation à une expression linguistique [se fait], selon les hypothèses montaguiennes, par l'intermédiaire d'une fonction interprétative préservant les structures d'algèbres entre le langage formel et l'univers d'interprétation.

Ceci signifie que le langage formel est défini comme un couple constitué d'un ensemble et d'une famille d'opérations sur cet ensemble. Les éléments de cet ensemble seront les expressions linguistiques (correctement formées), et les opérations seront les opérations constitutives de ces expressions, opérations que nous qualifierons de *syntaxiques*. Ceci signifie aussi, que, de façon analogue, l'univers d'interprétation est défini, lui aussi, comme un couple constitué d'un ensemble et d'une famille d'opérations sur cet ensemble, que nous qualifierons de *sémantiques*. Ceci signifie

encore qu'il y a une correspondance stricte entre les opérations syntaxiques et les opérations sémantiques, c'est-à-dire qu'à chaque opération syntaxique est associée une et une seule opération sémantique. Par contre à plusieurs opérations syntaxiques différentes, peut être associée une même opération sémantique.

La préservation des structures d'algèbre signifie enfin que si une expression linguistique ϕ est obtenue par application d'une opération syntaxique F à des expressions e_1, e_2, \dots, e_n , alors l'interprétation de ϕ sera obtenue par application de l'opération sémantique associée à F aux interprétations des expressions e_1, e_2, \dots, e_n [Chambreuil 98 p. 52-53].

C'est autour des Idées 1 à 3 et de leur statut méthodologique, et de l'homomorphisme algébrique visée entre LN et univers d'interprétation, que les interrogations de ce travail s'articulent. Remarquons que dans toutes ces citations il s'agit d'*expressions linguistiques*, ce qui pour nous laisse ouverte la possibilité d'une interprétation ambiguë : expressions graphiques ou expressions graphiques associées à une étiquette?

3 Les interrogations de ce travail en fonction de ses objectifs

Étant donné que le fil méthodologique conducteur de 5P est d'essayer de pratiquer la linguistique comme une science empirique, et que les Fonctions sémantiques de 5P doivent spécifier explicitement une représentation sémantique, les interrogations de ce travail par rapport à la Grammaire de Montague – définie de la manière étroite indiquée ci-dessus – s'articulent autour de deux questions :

- le statut d'hypothèses empiriques ou non des Idées 1 à 3 de la Section précédente ;
- l'exigence d'homomorphisme algébrique entre expressions linguistiques et univers d'interprétation.

Ces deux questions sont abordées successivement dans les deux Sections qui suivent. Montague a introduit les Idées 1 à 3 dans UG, où il les a pratiquées d'une manière assez particulière. C'est PTQ qui est plutôt le point de référence obligé lorsqu'on parle de Grammaire de Montague, pour ce qui est de l'utilisation du formalisme montaguien dans une description concrète.

Entre UG et PTQ les différences, nous l'avons remarqué dans notre litanie de difficultés, ne sont pas négligeables. Les trois Sections suivantes seront consacrées à étudier les deux questions précédentes dans le cadre de UG. La Section 4 sera consacrée à les étudier dans celui de PTQ.

3.1 L'interrogation sur l'empiricité

Dans les écrits de Chambreuil est constamment attribuée à Montague la formulation d'hypothèses. Montague aurait fait des hypothèses sur ceci et sur cela : sur la traduction des noms propres en formules lambda, sur l'objet de la sémantique, sur la structure d'algèbre de la langue naturelle... Comme il a été remarqué dans la Section 1.1, ceci est assez habituel

en linguistique et en particulier en linguistique formelle et /ou linguistique informatique, où on pourrait montrer facilement par cumulation de citations que n'importe quelle idée, en particulier si elle est présentée quelque peu habillée par des notations dans un formalisme quelconque, devient une hypothèse ou, pire encore, une théorie. Mais par rapport à Montague et à ses apports fondamentaux, il y a lieu de se demander : est-ce que ce qu'il a proposé ce sont bien des hypothèses propres à une science du réel?

Le rappel de la citation suivante de [Chambreuil 98 p. 46] permet de poser la question :

L'hypothèse fondamentale apparaît alors comme étant celle qui postule une architecture analogue pour les logiques, langages artificiels construits par les logiciens, et les langues naturelles.

Le problème est bien qu'une hypothèse d'une science du réel ne se postule pas : elle se propose, elle se teste, dans une optique popérienne elle n'est jamais vérifiée, on peut au maximum savoir à un moment donné qu'on n'a pas réussi à montrer qu'elle est fausse. Alors, qu'est-ce que Montague a fait ? Postuler une idée et tout adapter pour trouver dans la langue naturelle ce qu'il voulait trouver, ou proposer une hypothèse sur les langues naturelles dans le cadre d'une science du réel ? Ou est-ce encore quelque chose d'autre ?

S'il s'agit d'une vraie hypothèse d'une science du réel, il faut, au minimum, se donner un système d'observation du réel sur lequel elle porte permettant de la tester, parfois de manière indirecte, en testant les conséquences qu'on peut en déduire. Ce réel n'est pas préexistant ou immédiatement donné : il faut le convenir, l'explicitier. On ne demande donc pas à une science du réel qui traite d'un Réel en majuscule, mais d'un *réel-observé*, c'est-à-dire d'un Réel que l'on a convenu d'observer de telle ou telle manière. On a ainsi :

Convention d'observation (Réel) = réel-observé

Et il faut qu'il y ait un ou plusieurs réels observés, correspondant à différentes manières d'observer le Réel, et il faut que ce réel-observé puisse être utilisé pour tester les hypothèses empiriques et leurs conséquences qui sont censées porter sur lui.

En prenant la question de l'empiricité comme fil conducteur de fond pour interroger la Grammaire de Montague, nous adoptons un belvédère d'observation qui, explicitement, n'a pas été adopté dans les travaux de Chambreuil. Dans le premier des travaux qu'il a consacrés à la Grammaire de Montague [Chambreuil 89, p. 14], il indique que son étude :

[...] ne sera pas une réflexion sur la méthodologie de la grammaire de Montague (par exemple dans l'esprit des études de R.P. Botha), sur ces critères d'adéquation aux langues naturelles, à certains aspects de l'activité langagière.

Et ce choix a été conservé dans tous les travaux ultérieurs (cf. [Chambreuil 89], [Chambreuil & Pariente 90], [Chambreuil 91], [Chambreuil 98], [Chambreuil & al. 98]).

En adoptant le point de vue de l'empiricité, nous n'entendons pas nous consacrer à montrer que telle ou telle structure de l'anglais n'a pas été décrite dans UG ou dans PTQ, ou que telle description conduit à de mauvais résultats. Cet aspect des choses sera traité seulement de manière très marginale. L'interrogation sur l'empiricité porte sur une question de fond et non sur les limites – relativement étroites – de telle ou telle description particulière.

Les positions montaguennes concernant le pôle (i) de ses apports ont été synthétisées dans les Idées 1 à 3 présentées ci-dessous dans la Section 2.

Les Idées 2 et 3 sont le résultat d'un choix sur la manière d'observer le Réel pour obtenir le réel-observé. Bien entendu, ce choix sera caractérisé par celui qui le fait comme étant le bon choix, et la manière choisie pour observer le Réel sera caractérisée comme étant efficace, significative, éclairante, naturelle... Nous croyons qu'il ne revient qu'à l'histoire des sciences de pouvoir apprécier si, avec le temps et les apports d'une communauté scientifique, un choix initial est devenu ou non un bon choix.

Mais les choix initiaux peuvent aussi être, en quelque sorte, justifiés. L'Idée 2 est justifiée ainsi par Montague :

The basic aim of [...] syntax is to characterize the various syntactical categories, especially the set of declarative sentences. It is to be expected, then that the aim of syntax could be realized in many different ways, only some of which would provide a suitable basis for semantics. [...] I fail to see any great interest in syntax except as a preliminary to semantics (UG p. 223 note 2).

On dira ici que l'Idée 2 peut se caractériser comme un constat sur la non-observabilité de principe de certains objets – les objets syntaxiques – comme étant des objets susceptibles de tester, per se, les hypothèses sur les langues naturelles. Ou autrement dit, le réel observé ne peut pas être constitué par des arbres. Ce choix de la non-observabilité de principe des structures syntaxiques est le choix fait dans 5P, une trentaine d'années après Montague. Il a été largement inspiré par la prise de position, absolument pionnière, de Montague et il a été corroboré par ce qui s'est passé en linguistique, notamment dans le cadre chomskyen, depuis 1970, date de publication de UG.

L'Idée 3 n'est pas justifiée par Montague mais simplement posée. C'est une manière de définir ce qu'est – ou plutôt ce que doit être d'après lui – la sémantique d'une langue particulière. En tant qu'Idée posée, elle ne sera pas discutée *in extenso* ici. Mais on se limitera à signaler que les arguments apportés (par d'autres que lui) pour la justifier sont plutôt minces. On ne cite à continuation que ceux de [Halvorsen & Ladusaw 79, p. 185], mais des arguments analogues ont été exprimés ailleurs (cf. p. ex. [Partee 97 p. 9] et la citation qui y est faite de [Lewis 70]).

Linguistic semantics [contrairement à la *model theoretic semantics* proposée par Montague] as it has been practiced by generative and interpretive semanticists has been limited to translation from natural languages to various semi-formal languages (the language of semantic markers, the language of natural logic, etc.). This approach is of limited use. Since it deals only with translation,

it can at most explicate the meaning of expressions by rephrasing them in a different language. Understanding the meaning of an expression involves realizing a connection between the expression and extra-linguistic entities which it applies to. No such link can be modelled by a semantics which is only concerned with translation. Another limitation on the traditional forms of linguistic semantics stems from the lack of a formal characterization of the language in which the semantic representations are stated. This makes it impossible to explicate the meaning of an expression of, for example, English by translation into the language of semantic representations, since the latter is no better understood than English, and in fact is only interpretable (in the non-technical sense) thanks to our knowledge of English.

Si l'on peut adhérer sans réserves à l'absence de caractérisation formelle du langage dans lequel ont été ou sont exprimées certaines représentations sémantiques – absence qui est aussi notoire dans la spécification de la syntaxe de la langue naturelle dans PTQ, cf. Section 4 –, nous croyons qu'il est impossible d'établir, avec les termes utilisés dans la citation précédente, la dichotomie entre la « model theoretic semantics » proposée par Montague et les représentations non calculables.

UG et PTQ ne font autre chose que traduire d'un langage à un autre, même si dans un cas on nomme cette traduction *traduction* et dans l'autre *interprétation*. Cela non seulement parce que les expressions de la langue naturelle sont traduites en *form-LI* (ce sont les cas où on utilise *traduction*) mais, étant donné la dispensabilité de principe de LI, de manière beaucoup plus importante, parce que l'univers de dénotation n'est autre chose qu'un objet exprimé par un langage : ladite *interprétation* dans un univers suppose une *traduction* dans cet univers afin de vérifier si certaines conditions sont ou non satisfaites dans les représentations, en forme de langage, de cet univers. L'ensemble des « Individuals » n'est autre chose qu'un ensemble de symboles, les valeurs de vérité ne sont autre chose que deux symboles et ainsi de suite. C'est ainsi que la dernière phrase de la citation précédente peut s'appliquer, telle quelle, à la « model theoretic semantics » proposée par Montague. Comme exercice pratique dans ce cadre, qui est une bonne illustration de notre assertion, le lecteur est invité à lire les quasi trois pages de commentaires que [Dowty et al. p. 216-219] est contraint de consacrer à la form-LI censée donner l'interprétation *de dicto* de (i), form-LI reproduite ici dans (ii) :

- i John seeks a unicorn
- ii $\text{seek}'(j, \wedge \lambda Q \exists x [\text{unicorn}'(x) \wedge Q \{x\}])$

D'après [Dowty et al. p. 216] la formule en (ii) :

[...] asserts that John (the individual denoted by *j*) stands in the "seek-relation" to, as it were, the property of being a property that some unicorn has. We must try to understand exactly what the interpretation of this formula is. The interpretation is undoubtedly one of the most arcane aspects of the entire PTQ system, but also one of the most fascinating and ingenious.

Après quoi suivent, naturellement en anglais, plus de deux pages pour expliquer ce que doit être l'univers dans lequel cette formule doit être évaluée. Nous ne voyons pas en quoi strictement cela diffère méthodologiquement de l'exigence de l'anglais pour expliquer l'interprétation des représentations sémantiques non calculables.

Contrairement aux Idées 2 et 3, que l'on vient de commenter et qui sont au-delà ou à côté de la question de l'empiricité, l'Idée 1 peut – d'après nous, *peut* et non *doit* – être interprétée comme une affirmation portant sur des analogies ou identités profondes entre deux classes d'objets : les langues naturelles d'un côté et les langages artificiels des logiciens de l'autre. Ces analogies ou identités sont telles que les deux seraient passibles de traitement dans les termes de la « model theoretic semantics » proposée par Montague. Cf. par exemple la citation de [Kamp & Reyle 93 p. 22] :

These papers [c'est-à-dire "Montague's papers on natural language semantics"] showed once and for all that the model-theoretic approach towards natural language was viable. They also demonstrated that the model theories for natural and those for symbolic languages have a great deal in common.

Montague himself made a point of stressing the similarities and playing down the differences.

L'Idée 1 s'instancie chez Montague par l'exigence d'homomorphisme algébrique entre univers de dénotation et langage-LN. Si l'on assume que cet homomorphisme existe avec les langages artificiels des logiciens (« artificial languages of logicians », UG p. 222) et si l'on peut montrer qu'il existe aussi ou peut exister avec le langage-LN, on peut estimer que l'Idée 1 a été, en quelque sorte, testée et vérifiée.

Il nous a été impossible de trouver un seul indice permettant de dire que l'Idée 1 de Montague, exprimée dans les premières lignes de UG, devait ou pouvait être interprétée, selon ce qui est dit dans UG, comme hypothèse d'une science du réel. Nous essayerons de justifier cette affirmation dans la Section suivante, où on présente par ailleurs un tableau d'ensemble de la notion de compositionnalité.

3.2 L'interrogation sur la compositionnalité

Il est possible de détecter dans les écrits sur notre sujet au moins les cinq notions de compositionnalité suivantes :

- i *compositionnalité*^{aea} ou *compositionnalité*^{auberge espagnole autrefois}
- ii *compositionnalité*^{strad} ou *compositionnalité*^{systématicité traduction}
- iii *compositionnalité*^{pi} ou *compositionnalité*^{présentation intuitive}
- iv *compositionnalité*^{hu} ou *compositionnalité*^{homomorphisme unique}
- v *compositionnalité*^j ou *compositionnalité*^{janssen}

Nous avons fait déjà allusion à (i) *compositionnalité*^{aea} dans la Section 1.1. Nous affublons cette *compositionnalité* avec l'indice *auberge espagnole autrefois* car il s'agit typiquement d'un espèce de truisme que chacun peut interpréter à sa manière, le discuter aussi à sa manière, trouver des antécédents ici ou là, etc. Et cette situation résulte du fait que les points clefs d'un semblant de définition ne sont nullement établis. Cela apparaît

clairement dans [Partee & al. 90 p. 318] et c'est dit encore plus clairement dans [Partee 97 p. 20-21] :

The compositionality requirement [...] can be stated in plain language as follows :

The Principle of Compositionality :

The meaning of a complex expression is a function of the meanings of its parts and of the way they are syntactically combined.

Construed broadly and vaguely enough, the principle has sometimes seemed uncontroversial [...]

As the wording given above suggests, the exact import of the compositionality principle depends on how one makes precise the notions of *meaning*, of *part* and of *syntactic combination*, as well as on the class of functions permitted to instantiate the "is a function of" requirement.

L'existence de (i) est clairement illustré dans [Nazarenko 98]². Il se peut que (i) ait été une manière de poser les choses à un moment donné, mais nous n'avons rien de plus à dire à ce sujet. Dans la Section 6 nous pointons cependant quelques malheureux malentendus qui entourent l'interprétation de la Grammaire de Montague, largement attribuables à cette notion peu précise.

L'idée de (ii) compositionnalité^{strad} nous semble parfaitement claire et introduit une ligne de clivage importante dans la manière de concevoir la relation entre structures syntaxiques et représentations sémantiques, quelles que soient celles-ci. Elle n'est pas présentée avec le label que nous lui donnons ici, mais elle est clairement présente dans [Chambreuil & Pariente 90]. Ils soulignent avec raison que le fait même qu'il y ait une traduction explicite des suites-LN en form-LI est une caractéristique essentielle des Grammaires de Montague. L'explicitation de cette traduction est avantageusement comparée à la pratique de l'assignation directe des représentations logiques aux suites-LN, sans qu'aucun processus systématique pour ce faire ne soit spécifié. Cette observation revient à plusieurs reprises :

Concernant les rapports entre langue naturelle et logique des prédicats, on peut ici encore noter que le passage entre une expression de la langue naturelle et sa représentation sur le système formel de la logique des prédicats ne se fait pas par application d'un ensemble de règles formelles systématique. Il se fait par analyse de l'expression par l'utilisateur qui ensuite retranscrit son analyse dans les termes de la logique (p. 55).

Concernant cette formule de la logique des prédicats, on doit remarquer que sa mise en correspondance avec l'expression « Jean marche lentement dans un parc » repose de fait sur une capacité d'analyse sémantique de l'expression de la part de celui qui déterminera cette association. [...]

Ce que nous voulons souligner, c'est simplement le fait que n'est explicitement défini aucun mécanisme qui permettrait une association systématique entre les expressions du langage naturel et les expressions du langage de la logique des prédicats (p. 163-164).

Cette manière de voir les choses est donc explicitement comparée à l'attitude opposée de Montague, qui, lui, a cherché à se donner un outil de traduction :

En définissant dans UG une notion générale de traduction d'un langage dans un autre langage, en explicitant dans UG, puis dans PTQ, des exemplaires particuliers de cette notion générale, c'est un

² Pratiquement dans tous les articles du volume on revient sur cette absence de définition de compositionnalité.

tel mécanisme systématique d'association d'expressions de la logique intensionnelle aux expressions d'un langage naturel que cherche à caractériser le système de R. Montague (p. 164).

On ne peut qu'adopter la compositionnalité^{strad} de (ii) comme une exigence nécessaire sur l'association des représentations sémantiques aux expressions linguistiques; on reprend ce point dans la Section 8. Il reste évidemment à instancier cette exigence de base³.

La compositionnalité^{Pi} (cf. (iii) compositionnalité^{présentation intuitive}) est l'idée de la compositionnalité que l'on peut légitimement se faire à partir des présentations intuitives qui sont données dans des exemples volontairement simples et choisis en tant que tels. Nous en citons plusieurs, et des meilleurs auteurs – Chambreuil et Dowty – qui ont fait des présentations rigoureuses des Grammaires de Montague. Selon [Dowty el 81] :

It is tempting, however, to regard the semantic value given to an expression which is not a sentence as "the meaning" of that expression, and if we allow ourselves to do as a first approximation, we can state Frege's Principle in the form in which it is most often quoted : The meaning of the whole is a function of the meanings of the parts and their mode of combination. [...] Note, for example, that "the parts" referred to in the statement of Frege's Principle must be syntactic constituents of the expression in question (p. 8-9).

L'interprétation légitime de ce texte est que les « syntactic constituents » ont comme suite terminale des suites de « linguistic forms » (c'est-à-dire des suites d'objets qu'on ne peut manipuler que par leur aspect graphique), que le « whole » est aussi une suite, et que les suites terminales des « syntactic constituents » sont des suites (dans le sens restreint de la Section 1.1) à l'intérieur du « whole », ce « whole » étant une phrase complète. Ce qui veut dire que si l'on a une suite s telle que

$$s = [a, b, c, d, e]$$

l'interprétation sémantique de s viendra de l'interprétation des constituants qui, eux-mêmes, vont avoir comme des suites terminales des sous-suites de s . Ainsi, l'interprétation de s doit venir de constituants qui ont comme suites terminales $[a]$, ou $[a, b, c]$, ou $[d, e]$, ... etc. Ce serait justement le rôle de la syntaxe de déterminer au mieux ces sous-suites pour leur donner des valeurs sémantiques à partir desquelles calculer la valeur sémantique de s .

Remarquons que c'est exactement comme cela que les choses se passent, autant que l'on sache, dans tout langage de la logique (logique propositionnelle, logique de premier ordre avec ou sans identité, logique intensionnelle de Montague, DRT à la Kamp). On définit des classes élémentaires d'entités par leurs caractéristiques graphiques notationnelles : on pointe ainsi les variables de tel type en disant qu'elles sont notées avec tel type de caractères, on pointe les symboles syntaxiques, les opérateurs, les quantificateurs, etc. simplement en montrant le graphisme qu'on utilise pour les noter. En plus de ce repérage graphique des entités sur lesquelles on va travailler, on se donne les conditions de bonne formation en fonction de règles sur la concaténation des suites adjacentes bien formées. À

³ [Poirier & al 98] et [Amsili & al. 98] sont, parmi d'autres, des propositions pour ce faire.

chaque constituant syntaxique on va donner une interprétation sémantique. De telle manière que l'interprétation donnée à la citation précédente n'est pas tendancieuse : c'est ce que l'on fait en logique. Comme exemple tiré du même [Dowty & al. 81], voir le passage suivant :

We then state *semantic rules*, whose job is to determine the semantic value of each larger constituent in terms of the semantic values of its components. In other words, semantic values are assigned to successively more inclusive constituents of the sentence until finally the semantic value of the entire sentence has been determined (p. 16).

Cette interprétation de la compositionnalité coïncide entièrement avec celle qui est donnée dans, p. ex., [Chambreuil 98, p. 53] :

Ainsi, selon cette hypothèse, appliquée récursivement à une expression et à ses constituants, l'interprétation d'une expression complexe est totalement déterminée par l'interprétation des expressions élémentaires qui la constituent et la séquence des opérations syntaxiques mises en jeu dans sa construction. C'est dans ce sens qu'une telle interprétation sera dite *compositionnelle*.

Toutes les présentations intuitives dans les travaux de Chambreuil admettent l'interprétation donnée ci-dessus. Considérons la suivante, déjà citée ci-dessus à l'intérieur d'un texte plus étendu :

Pour une langue naturelle, cela [la compositionnalité] se traduira, en particulier, par les éléments suivants. Par exemple, une entité de l'univers d'interprétation sera associée à « Jacques », une autre entité sera associée à « parler ». Une opération (syntaxique) permettra de construire « Jacques parle ». À cette opération correspondra une opération sur les entrées associées respectivement à « Jacques » et à « parler » qui (hors problème de temporalité) appliquée à ces entités aura pour résultat l'entité de l'univers d'interprétation associée à « Jacques parle » (Chambreuil 98, p. 45-46).

Et pour conclure sur ce point, observons l'orientation générale de la Grammaire de Montague explicitée dans [Chambreuil & Pariente 90] citée par la suite :

(1) Le langage formel qui sera visé correspondra cette fois, directement, à un langage naturel.

(2) Les expressions élémentaires seront des expressions élémentaires pour le langage naturel (ou considérées comme telles).

Ainsi l'expression langagière « Jean marche lentement dans le parc » ne sera plus associée à une expression élémentaire [de la logique des propositions] comme c'était le cas en logique propositionnelle.

Elle ne sera pas non plus associée à une expression [de la logique des prédicats]. L'expression langagière sera obtenue elle-même comme expression du langage formel, construite à partir des expressions élémentaires : « Jean », « marcher », « lentement », « dans », « parc », et aussi « un » (p. 64).

Nous ne croyons pas que l'interprétation – légitime – que l'on a le droit de tirer de ces présentations intuitives soit la bonne interprétation de la compositionnalité selon la Grammaire de Montague. Il n'y a rien dans la Grammaire de Montague qui exige que l'interprétation sémantique d'une suite *s* viendra exclusivement de l'interprétation des

constituants qui, eux-mêmes, vont avoir comme suites terminales des sous-suites de *s*. Rien dans UG comme nous le verrons ci-dessous dans la Section 3.2.1, ni rien dans la pratique concrète de PTQ, qui en est extraordinairement éloignée comme nous le verrons dans la Section 4.

La compositionnalité de UG est la compositionnalité^{hu} (iv) ou compositionnalité^{homomorphisme unique}. Nous l'aborderons en détail dans la Section suivante.

La compositionnalité^j (v) (compositionnalité^{janssen}) a été formellement définie dans [Janssen 97]. Très proche de la précédente elle ne sera pas analysée exhaustivement, mais la position méthodologique de [Janssen 97], à laquelle nous ne souscrivons pas, sera reprise dans les Sections 7 et 8.

3.2.1 La compositionnalité dans UG

Dans cette Section on va s'occuper de la compositionnalité^{hu} (iv), désormais labellisée homomorphisme-u, utilisée dans UG.

On trouve dans UG :

- i les Idées 1 à 3 rapportées dans la Section 2 et dont l'empiricité a été déjà partiellement discutée dans la Section 3.1 ;
- ii un cadre formel où sont définies formellement les notions qui vont permettre d'associer un langage à l'univers de dénotation, soit directement, soit en passant par la traduction à un autre langage ;
- iii la formulation de LI (logique intensionnelle) ;
- iv l'utilisation de (ii) dans la description d'un fragment de l'anglais.

Le point (i) a été déjà partiellement discuté dans la Section 3.1. Cette discussion sera complétée ici par l'examen de (ii). Nous n'avons rien à dire sur (iii). La description de (iv) n'est pas le point de référence sur l'utilisation de la Grammaire de Montague dans la description d'un fragment de l'anglais. Par ailleurs, il faut remarquer que dans la description de (iv), Montague continue à utiliser sévèrement ce que Partee appelle la *mise au pas* (anglais *regimentation*, cf. [Partee 97 p. 12]).

En effet, dans la philosophie du langage, et ceci déjà depuis Port Royal⁴, il y a une tradition qui, brutalement dite, consiste dans ceci : on ne philosophe pas sur ce que le langage est dans ce qu'il a d'observable, ou sur quelque chose qui doit être intégré à ce que le langage a d'observable, mais sur un objet qui a été modifié ou déformé quasiment *ad libitum*. Non

⁴ [Pariente 85 p. 129-130] observe que « quatre pages et demi disposent de la syntaxe [...] Quant à la syntaxe de régime[...] à quelques exceptions près, elle varie de langue à langue : de ce fait elle n'a pas sa place dans une Grammaire générale. Tels sont les arguments explicites de Port-Royal ».

que les hypothèses sur le réel-observé doivent porter exclusivement sur ce qui est parfaitement et directement observable exclusivement. Mais oui au besoin impératif de spécifier le lien, les limites du décalage, entre toute idéalisation de l'objet à étudier et tout ce qui est intersubjectivement observable dans l'objet. Port Royal a procédé comme si la syntaxe dite aujourd'hui de surface et la morphologie étaient des curiosités bizarres dont on pouvait se passer pour ainsi trouver la grammaire générale. Quine s'est donné un anglais-Quine (cf. [Bès 85]) pour traiter les lectures *de dicto*. Cette tradition est bien caractérisée par [Partee 97 p. 11-12] :

Much insight into the semantic content of natural language expressions was achieved in these studies [études sur la philosophie du langage non cités in extenso par Partee], but relatively little progress was made on systematically relating semantic content to syntactic structure. For those natural language constructions where the semantically relevant syntactic structure was no perfectly straightforward (which were many), the strategy was "regimentation": the invention and analysis of formal languages which contained syntactically "transparent" analogs of the constructions of interest, languages which met the logician's criterion for being "logically perfect" : unambiguous and describable with a unique semantic interpretation rule for each syntactic formation rule.

Or la mise au pas de l'anglais subsiste dans UG. Pour ne citer qu'un seul exemple, on ne représente ici que celui évoqué dans [Partee 97 p. 52]. Dans UG on ne décrit pas l'anglais de la première phrase qui suit, mais l'anglais mis au pas de la deuxième.

- i Every man loves a woman which loves him.
- ii Every man is an entity such that it loves a woman that she loves it.

Il nous reste ainsi à étudier le point (ii) de UG, c'est-à-dire le cadre formel où sont définies formellement les notions qui vont permettre d'associer un langage à l'univers de dénotation, soit directement, soit en passant par la traduction à un autre langage. C'est l'examen quelque peu détaillé de ce cadre qui va permettre de prendre position par rapport au statut méthodologique de l'Idée 1 de Montague rapportée dans la Section 2. Mais on peut déjà avancer que ce cadre n'est pas une théorie dans le sens un peu précis avec lequel nous avons utilisé ce terme selon ce qui a été stipulé dans la Section 1.1.

Le système de définitions de UG, en ce qui concerne notre problème, s'organise dans les treize points qui suivent.

- i Définition d'algèbre.
- ii Définition de la classe K des opérations polynomiales sur un algèbre telle qu'elle est définie dans (i).
- iii Définition de la similarité entre algèbres.
- iv Étant donné Alg1 et Alg2 -un couple d'algèbres similaires (cf. Déf (iii)), définition de
 - l'homomorphisme *de* (anglais *from*) A1 *vers* (anglais *into*) A2
 - l'homomorphisme *de* A1 *sur* (anglais *to*) A2.
- v Définition de « langage non ambigu » (cf. le caveat terminologique introduit dans la Section 1.1).

- vi Définition de la famille C de catégories syntaxiques générées par un « langage non ambigu ».
- vii Définition de « langage ».
- viii Interprétation d'un « langage ».
- ix Assignation de signification à un « langage », qui est un homomorphisme unique
[La définition de (ix) introduit des restrictions à la définition plus générale de (viii); nous dirons ici *homomorphisme-u d'assignation de signification*].
- x Définitions concernant la signification.
- xi Définition d'une base de traduction.
- xii Définition d'une fonction de traduction, qui est un homomorphisme unique
[nous dirons ici *homomorphisme-u de traduction*].
- xiii Définition d'interprétation induite d'un « langage » par un autre « langage ».

Le point culminant de cette série de définitions est bien entendu (xiii). C'est dans (xiii) qu'est établie la dispensabilité de principe de la LI dans l'interprétation du « langage »-LN dans un univers de dénotation, tout en conservant la possibilité de cette interprétation. Pour aboutir à (xiii), UG présente quatre « Remarks »⁵ qui sont en fait des conjectures sur des théorèmes. La mécanique d'ensemble de la démonstration formelle repose principalement sur :

- la définition (ii), qui va permettre de définir, étant l'algèbre de LI, ce que nous appellerons ici *algèbre dérivée* à partir des opérations de LI (nous notons Op^{LI} les opérations de LI, et Ad^{LI} l'algèbre dérivée à partir de Op^{LI}) ;
- la Remarque 1 (p. 225), qui établit que si h est un homomorphisme sur Op^{LI} , $h(Ad^{LI})$ va nous donner l'algèbre unique A' telle que h est un homomorphisme de Ad^{LI} sur A' ;
- l'homomorphisme-u d'assignation de signification de Ad^{LI} sur l'algèbre unique A' ;
- l'homomorphisme-u de traduction du « langage non ambigu » dans Ad^{LI} ;
- la relation R , qui donne $R(\text{« langage non ambigu »}) = \text{« langage ambigu »}$.

L'homomorphisme-u d'assignation de signification du « langage non ambigu » dans A' est le résultat de la composition de l'homomorphisme-u de traduction du « langage non ambigu » dans Ad^{LI} avec l'homomorphisme-u d'assignation de signification de Ad^{LI} sur l'algèbre unique A' (cf. UG Remarque 4 p. 233). C'est grâce à cette composition et à la relation R qu'il sera possible d'associer aux suites du « langage » une signification par un chemin plutôt indirect qui passe par le « langage non ambigu », l'algèbre dérivée de LI (Ad^{LI}), elle-même obtenue par des opérations polynomiales sur les opérations de LI (Op^{LI}), la traduction du « langage non ambigu » sur cette algèbre dérivée de LI et l'interprétation de cette algèbre dérivée dans l'algèbre des significations (A'), ces significations étant bien entendu celles qui sont requises par LI.

⁵ Ces Remarques ne sont pas numérotées dans UG; nous les identifions selon la page où elles ont été exprimées.

La présentation précédente résume le cadre formel où est posée l'Idée 1 de la Section 2, qui est celle qu'il nous intéresse d'interroger sous l'angle de son empiricité éventuelle, afin de pouvoir lui donner un statut méthodologique plus clair.

On va organiser la présentation du résultat de l'analyse de UG dans les Ob (Observations) 1 à 3 qui suivent. Ob1 et Ob2 reprennent la question de la « compositionnalité » chez Montague et, à leur suite, on prend position sur le statut de l'Idée 1. L'Ob 3 s'attaque à l'un des problèmes qui semblent cruciaux dans l'application aux langues naturelles de l'homomorphisme- α d'assignation de signification et de l'homomorphisme- α de traduction, le problème de l'assignation catégorielle ambiguë possible aux expressions élémentaires d'un « langage non ambigu ».

Ob. 1 Montague n'a pas utilisé *compositionnalité* ou ses dérivés comme des notions techniques. Il fait référence au truisme de la compositionnalité au cours des discussions sur les définitions de (x) concernant la signification, p. ex. lorsqu'il introduit la différence intuitive entre *meaning* et *sens* :

The intuitive distinction is this : meanings are those entities that serve as interpretations of expressions (and hence, if the interpretation of a compound is always to be a function of the interpretations of its components, cannot be identified with functions of possible worlds alone), while *senses* are those intensional entities that are some times *denoted* by expressions [UG, p. 228].

C'est toujours dans (x) qu'il introduit l'expression *Fregean interpretation*, alors qu'il a déjà introduit les définitions formelles de (viii) et (ix). Avec *Fregean interpretation* on n'a pas besoin de comprendre que Montague attribue à Frege la « compositionnalité »; plutôt on peut ou doit comprendre qu'avec les notions qui lui sont propres de (viii) et de (ix), plus la manière de concevoir les *meanings* qu'il pose dans (x), il est capable de rendre compte des observations de Frege sur *Sinn* et *Bedeutung*, [Frege 1892] étant le seul travail cité de Frege dans UG.

Les notions formelles de (viii) et (ix), qui sont parmi les notions clés pour aboutir à (xiii) sont complètement indépendantes des définitions de (x), dans ce sens que ce qui est défini dans (x) n'est exigé en rien pour que les notions de (viii) et (ix) soient susceptibles d'être appliquées. Les entités de (x) peuvent être, formellement, des entités arbitraires, pourvu qu'elles soient organisées sous forme d'algèbre. Comme [Halvorsen & Ladusaw 79 p. 201] l'observent :

Section 3 [d'UG] is concerned with how meanings are associated with expressions of a language ["language", dans notre système de notation] [...] by means of interpretations. Montague does not at this point specify what kind of things meanings have to be. The possibility is left open for meanings to be whatever type of object one might like, the decision presumably being dependent on the preference of the analyst and the special requirements of the problem dealt with.

De même, les notions formelles de (ix) et (xii), qui sont aussi parmi les notions-clés pour aboutir à (xiii) sont complètement indépendantes du « langage non ambigu » particulier qui sera l'aboutissement de la traduction, pourvu que celui-ci soit susceptible d'être organisé en algèbre, dont on peut tirer une algèbre dérivée, ce qui est garanti par la définition même de « langage non ambigu ».

Ob. 2 Dans cette Observation nous allons enregistrer une série d'inexistences qui pourraient se résumer dans la formule générale suivante : tout ce qu'il y a de définitions formelles dans UG se limite aux points (i) à (xiii) ci-dessus. Pour expliciter les inexistences clés, il est nécessaire d'introduire au préalable quelques conventions de notation.

On connaît le quintuplet qui définit un « langage non ambigu » :

$$\langle A, F_{\gamma}, X_{\delta}, S, \delta_0 \rangle_{\gamma \in \Gamma, \delta \in \Delta}$$

Nous allons noter $L-M$ (avec M de Montague, naturellement), ce qui est noté A dans le quintuplet, et $SFL \sim A-M$ tout ce qui est à droite de A dans le quintuplet, à savoir :

$$\langle F_{\gamma}, X_{\delta}, S, \delta_0 \rangle_{\gamma \in \Gamma, \delta \in \Delta}$$

On a ainsi : « langage non ambigu » = $\langle L-M, SFL \sim A-M \rangle$

Par ailleurs, nous allons noter Op ce qui dans le quintuplet est noté F_{γ} et nous allons noter CBF (Couples de Bonne Formation) ce qui dans le quintuplet est noté S .

On sait que :

- $L-M$ est clos pour Op ;
- X_{δ} est un ensemble des suites élémentaires -au sens strict donné à suite dans Section 1.1, c'est-à-dire un objet graphique non associé à rien d'autre- de longueur 1 qui sont dans un ensemble indexé par δ : c'est-à-dire que c'est l'ensemble qui est indexé et non les suites élémentaires qui appartiennent à cet ensemble ;
- $\cup_{\delta \in \Delta} X_{\delta} \subset L-M$;
- « langage » = $R(\langle L-M, SFL \sim A-M \rangle)$.

À partir de là, on peut noter les inexistences suivantes de spécification de contraintes :

- sur les entités de Δ ,
- sur Op ,
- sur CBF (qui se définit à partir de Op et de Δ),
- sur R .

Ces inexistences ont déjà été, en grande partie, signalées. On présente par la suite un certain nombre de citations qui en témoignent.

Dans l'avant-dernier chapitre de [Dowty & al. 81], après présentation du système de PTQ, au fond d'une note, on trouve l'assertion suivante :

In Montague's system, the way of combining one expression with another and the syntactic informations of the two expressions that result from this process are practically unlimited, though of course the syntactic operation associated with each syntactic rule must specify what is to be done [...]. We know of no mathematical results about the generative capacity of Montague's syntax (note 1 de p. 248).

[Chambreuil & Pariente 90 p. 98-103] confirme ces inexistentances :

À travers cette discussion et ce renvoi aux définitions générales d'un langage dans la théorie de UG, ce qu'on doit noter, c'est la distance entre la caractérisation théorique générale et une caractérisation particulière, que ce soit pour définir le langage d'une logique ou celui d'une langue naturelle.

UG caractérise un langage comme une structure à huit éléments s'appuyant principalement sur la combinaison entre opérations [Op], règles d'attribution de catégorie [CBF], relation R. Mais sur ces trois types d'éléments fondamentaux, seules les règles [CBF] ont une forme contrainte par la caractérisation théorique. Par contre, aucune contrainte n'est formulée sur la forme même des opérations ou de la relation R. Quand on sait la généralité des notions d'opération et de relation, on mesure l'immense liberté qui est ainsi laissée à toute caractérisation particulière d'un langage. Une autre conséquence de cette absence de contrainte sera que de l'échec d'une caractérisation particulière à capter des propriétés d'un langage, on ne pourra conclure à la non-pertinence de la structure générale (p. 101).

Comme commentaire à cette citation, nous notons que les seules contraintes de *CBF* consistent à incorporer dans chaque *cbf* \in *CBF* une *op* \in *Op*, et à attribuer des $\delta \in \Delta$ aux arguments d'entrée et à la sortie de *CBF*. Comme aussi bien *Op* et Δ ne sont contraints par rien, on ne voit pas comment ne pas conclure que les *CBF* ne le sont pas non plus.

Bien entendu, *aucune contrainte* veut dire, strictement, aucune. Comme [Partee 97]⁶ le remarque justement :

In such a rule, F_i is a syntactic operation [*op*]; it may be as simple as concatenation or, as far as the requirements of UG are concerned, arbitrarily complex and not even necessarily computable [Partee 97 p. 21].

[...] Montague put no constraints on the syntactic operations [Op] that could be used in a syntactic rule [CBF], not even a constraint of computability [Partee 97 p. 48].

Les conséquences de tout cela ne peuvent être que claires : étant donné n'importe quelles significations formulables sous forme d'algèbre, n'importe quel « langage non ambigu » pourra être traduit sur ces significations, et toute langue naturelle formulée comme un « langage » pourra être associée à ce « langage non ambigu ». Il n'y a rien dans le cadre

⁶ Il est assez stupéfiant de lire dans un même travail et sous la responsabilité d'une même plume les affirmations rapportées ci-dessus dans le texte et la suivante : « [...] Montague's precise version of it ["The Principle of Compositionality"] places strong constraints on admissible systems of syntax and semantics » [Partee 97 p. 21].

formel de UG, synthétisé par les points (i) à (xiii), qui l'empêche ou qui soit utilisable dans l'état pour discriminer les langues naturelles des « langages artificiels des logiciens ».

Si on l'observe depuis l'interrogation sur l'empiricité, l'Idée 1 de Montague, telle qu'elle est présentée dans UG, est une non-hypothèse. Cela, encore une fois, a déjà été remarqué :

Together the results of Janssen and Zadrozny illustrate that without constraints on syntax and semantics, there are no counterexamples to compositionality. This gives the pleasant feeling that a compositional treatment is somehow always possible [Janssen 97 p. 456-457].

Dans UG aucune restriction sur la syntaxe et la sémantique n'est proposée. La compositionnalité^{hu} de UG est un cadre vide et on voit mal qu'il le reste tant que l'on ne posera pas des restrictions sur les « langages non ambigus » qui pourraient être associés aux langages-LN. Dans la Section 4 nous allons analyser les *Op* et les *CBM* effectifs utilisés dans PTQ pour décrire un fragment de l'anglais : l'analyse de PTQ devrait compléter la vision du problème, analysé dans la présente Section en ce qui concerne UG.

Ob. 3 Étant donné les définitions de UG, il est impossible, dans la description des LN, d'assigner une suite élémentaire d'un « langage non ambigu » de manière ambiguë à des ensembles indexés différemment afin obtenir des assignations de significations différentes, car, si on le faisait, on obtiendrait des résultats inadéquats. Ce point est repris dans l'Annexe *Grammaire de Montague et ambiguïté*.

Cela veut dire que, par exemple, si on a une suite élémentaire σ , elle ne peut pas appartenir à deux ensembles différents indexés par δ et δ' , avec δ et $\delta' \in \Delta$, $\delta \neq \delta'$, si l'on veut obtenir des significations différentes pour chacune des assignations catégorielles. Selon cette Obs. 3, l'impossibilité d'assignation ambiguë des suites élémentaires d'un « langage non ambigu » ne vient donc pas des définitions formelles sur ceux-ci, mais des résultats inadéquats que l'on obtiendrait⁷.

Plus spécifiquement, la source de l'impossibilité d'assignation ambiguë doit être trouvée soit dans l'interprétation du « langage non ambigu » (cf. Définitions (ix) de la Section 3.2.1) soit dans sa traduction (cf. Définitions (xii) de la même Section). Après présentation de quelques conventions de notation, nous examinons dans le détail les conséquences qui relèvent de l'interprétation, sur lesquelles peuvent se calquer celles relevant de la traduction, et ensuite, de manière plus abrégée, nous examinons celles-ci. Cela devrait nous

⁷ Ce sont les conditions 4 et 5 de UG p. 225 qui garantissent la non-ambiguïté, et, strictement, elles n'interdisent pas l'assignation catégorielle ambiguë des suites élémentaires dans un « langage non ambigu ». À noter que par « assignation catégorielle ambiguë des suites élémentaires » nous comprenons, comme dans UG, l'appartenance d'une même suite, expression graphique nue, à des ensembles indexés par des étiquettes catégorielles différentes, et non le fait d'utiliser des expressions associées en tant que telles à des étiquettes catégorielles. Nous n'avons donc pas *chien_{nom}* mais *chien* appartenant à un ensemble {chien, ...}_{nom}.

permettre de prendre position par rapport au problème évoqué ci-dessus Section 1.1 *in fine*, tel qu'il est présenté dans les ouvrages d'introduction à la Grammaire de Montague⁸.

Selon les conventions établies dans la formulation de Ob 2, le quintuplet :

$$\langle A, F\gamma, X\delta, S, \delta_0 \rangle_{\gamma \in \Gamma, \delta \in \Delta}$$

qui définit un « langage non ambigu », peut s'écrire par l'une ou l'autre des formules suivantes :

$$\langle L-M, Op, X\delta, CBF, \delta_0 \rangle_{\delta \in \Delta}$$

$$\langle L-M, SFL \sim A-M \rangle \text{ avec } SL \sim A-M = \langle Op, X\delta, CBF, \delta_0 \rangle_{\delta \in \Delta}$$

On sait que $L-M$ est clos pour Op . Le couple $\langle L-M, Op \rangle$ spécifie donc une algèbre, qui sera ici notée $\langle Al-M \rangle$ (avec M toujours mnémotechnique de *Montague*). De telle manière que nous avons encore deux autres manières d'écrire la formule qui note un « langage non ambigu » :

$$\langle \langle L-M, Op \rangle x \rangle$$

$$\langle Al-M x \rangle$$

$$\text{où } x \text{ s'instancie par } \langle X\delta, CBF, \delta_0 \rangle_{\delta \in \Delta}$$

Lorsqu'il s'agira du « langage non ambigu » de LI, les deux dernières formules deviendront respectivement

$$\langle \langle LI, OpLI \rangle x \rangle$$

$$\langle AlLI, x \rangle$$

avec, en plus, pour noter l'algèbre dérivée qui résulte de l'application des opérations polynomiales sur $OpLI$: $\langle AdLI, x \rangle$.

Nous avons ainsi plusieurs notations qui permettent de faire référence aux différents éléments qui intègrent le quintuplet définitoire d'un « langage non ambigu ».

La série de définitions qui culminent avec celle sur l'homomorphisme- u d'assignation de signification (cf. les définitions (ix) de Section 3.2.1) rendent impossible une assignation catégorielle ambiguë aux suites élémentaires d'un « langage ambigu » pour les raisons suivantes.

Étant donné :

- $\langle \langle L-M, Op \rangle x \rangle$

- « langage » = $R(\langle \langle L-M, Op \rangle x \rangle)$

⁸ Nos conclusions vont porter sur l'assignation catégorielle ambiguë des suites élémentaires, alors que dans [Chambreuil & al 98] on envisage la possibilité d'assignation ambiguë pour les suites qui ne sont pas élémentaires, c'est-à-dire qui résultent d'une opération, mais nous croyons que ce qui sera dit ici est extrapolable à ce dernier type de suites. Nous rapelons que par *suite* nous entendons des objets purement graphiques, non associées donc à aucune autre information.

l'interprétation de « langage » se définit moyennant le système interprétatif (« system » dans UG p. 227) $\langle B, Op', f \rangle$ qui détermine la fonction g – homomorphisme-u d'assignation de signification – définie par les éléments suivants :

- $\langle B, Op', f \rangle$
- $\cup_{\delta \in \Delta} X_{\delta} \subset L-M$
- $e \in \cup_{\delta \in \Delta} X_{\delta}$
- $f(e) = e'$ tel que $e' \in B$
- $f \subset g$

Rappel : les entités de B sont des entités arbitraires (cf. Ob 1 in fine).

À partir de g il est possible de définir l'assignation de signification (« meaning assignment » dans UG p. 227) à « langage ». Selon les définitions de (vi) de Section 3.2.1, ci-avant, pour tout $\delta \in \Delta$, $X_{\delta} \subset L-M$ et $X_{\delta} \subseteq C_{\delta}$, $C_{\delta} \in C$. Selon les définitions de (ix) de Section 3.2.1 :

- étant donné $\sigma' \in ME_L$ (ME_L : ensemble de « meaningful expressions » de « langage »)
- si
- $\sigma \in \cup_{\delta \in \Delta} C_{\delta}$
- $\sigma R \sigma'$
- $g(\sigma) = a$
- alors σ' *signifie* a (« means » dans UG p. 227)

Autrement dit, la suite σ' bien formée (appartenant aux « meaningful expressions ») reçoit sa signification (« meaning ») via la suite σ (ce seraient des suites $\sigma^1, \sigma^2, \dots, \sigma^n$ dans le cas des ambiguïtés) spécifiée par la famille C de catégories syntaxiques du « langage non ambigu » à laquelle est rattachée par R la suite σ' , R étant ce qui permet de relier « langage non ambigu » et « langage ».

Dans ce cadre de définitions, il s'agit de montrer ce qui se produirait si une suite élémentaire du « langage non ambigu » serait catégoriellement ambiguë. Supposons cette situation d'ambiguïté. On aurait, p. ex. :

$$\begin{aligned} X_{\delta} &= \{\sigma, \dots\} \\ X_{\delta'} &= \{\sigma, \dots\} \end{aligned}$$

avec δ et $\delta' \in \Delta$, $\delta \neq \delta'$.

Dans $\cup_{\delta \in \Delta} X_{\delta}$ – qui est un ensemble – on aurait un seul σ . Il est important de noter que l'on a σ dans $\cup_{\delta \in \Delta} X_{\delta}$ et non σ_{δ} et $\sigma_{\delta'}$, les indices δ et δ' portant sur les ensembles X et non sur les éléments de ces ensembles. On sait par ailleurs que $X_{\delta} \subseteq C_{\delta}$ et que $X_{\delta'} \subseteq C_{\delta'}$. Pour les mêmes raisons que précédemment, dans $\cup_{\delta \in \Delta} C_{\delta}$ il y a un seul σ .

Par définition de f , qui est une fonction, on aurait comme évaluation de $f(\sigma)$ un seul élément dans B du système interprétatif $\langle B, Op', f \rangle$. Or, supposons $\sigma' \in ME_L$. Pour que la

suite σ' soit sémantiquement ambiguë (« semantically ambiguous » de UG p. 227) il faut qu'il y ait plus d'un élément – p. ex. σ^1 et σ^2 dans $\cup_{\delta \in \Delta} C\delta$ tels que :

- $\sigma^1 R \sigma'$
- $\sigma^2 R \sigma'$

pour pouvoir obtenir :

- $f(\sigma^1) = a^1$
- $f(\sigma^2) = a^2$

Mais, comme on vient de le voir, dans $\cup_{\delta \in \Delta} C\delta$ il n'y a pas deux éléments σ^1 et σ^2 , mais un seul σ , de telle manière que l'assignation catégorielle ambiguë de σ dans le « langage knon ambigu » ne provoque pas une assignation ambiguë de signification de σ , alors que l'on a $\sigma \in X\delta$ et $\sigma \in X\delta'$. Dans ce cas, pour pouvoir obtenir formellement une assignation de signification au seul σ , on doit compléter le système ajoutant une stipulation ad hoc qui indique selon laquelle des pluralités des assignations catégorielles existantes pour σ , l'assignation de signification doit être faite.

À partir de ce qui précède et pour des raisons formelles tout à fait analogues, il est possible d'expliciter les résultats inadéquats dans la traduction d'un « langage non ambigu » selon les définitions (xii) de la Section 3.2.1 qui introduisent l'homomorphisme- u de traduction. On fera donc ici cette explicitation en raccourci.

Soit les deux « langages non ambigus »

- $\langle \langle L-M, Op \rangle x \rangle$
- $\langle \langle L-M', Op' \rangle x' \rangle$

Une base de traduction (« translation base » de UG p. 232) est un triplet $\langle g, Opd', j \rangle$ où

- g est une fonction de Δ de $\langle \langle L-M, Op \rangle x \rangle$ à Δ' de $\langle \langle L-M', Op' \rangle x' \rangle$;
- Opd' est obtenu par application des opérations polynomiales à Op' ;
- j est une fonction aux effets tout à fait analogues aux effets de f dans le système interprétatif $\langle B, Op', f \rangle$. Le domaine de j est $\cup_{\delta \in \Delta} X\delta$, c'est-à-dire le même domaine que celui de f dans le système interprétatif $\langle B, Op', f \rangle$, le co-domaine de j étant la catégorie $C'g(\delta)$ de la famille de catégories C' générées par $\langle \langle L-M', Op' \rangle x' \rangle$.

Il apparaît ainsi clairement comment, pour la traduction, on retrouve les mêmes effets pervers que pour l'interprétation. À partir de l'utilisation de j toute la chaîne de l'interprétation induite va donner une seule interprétation induite à e' :

Étant donné

- $e \in \cup_{\delta \in \Delta} X\delta$
- $j(e) = e'$ tel que $e' \in C'g(\delta)$

Les résultats que l'on obtient aussi bien dans l'interprétation que dans la traduction des « langages non ambigus » lorsque des suites élémentaires ont été spécifiées de manière catégoriellement ambiguë, à condition d'ajouter au système d'ensemble des stipulations ad hoc, sont formellement valides, mais ils sont inadéquats si l'on souhaite assigner des significations différentes – soit dans des interprétations « directes » soit dans des interprétations induites via la traduction – à des suites bien formées du langage à décrire.

Supposons les langages-LN et supposons que la traduction des « langages non ambigus » correspondants s'effectuent vers $\langle \langle LLI, OpLI \rangle xLI \rangle$, le « langage non ambigu » de LI. Or, si l'on accepte l'Observation généralisée suivante sur les langues naturelles :

Dans toute LN il y a des suites élémentaires catégoriellement ambiguës sur le plan morphosyntaxique telles qu'elles doivent recevoir, en corrélat à leur ambiguïté syntaxique, des assignations différentes de signification.

on doit conclure que, pour pouvoir rendre compte des instances concrètes de cette Observation généralisée, l'assignation catégorielle ambiguë conduit, dans le système UG, à des résultats inadéquats.

Comme exemple concret, soit la suite élémentaire de l'anglais *judge* (le raisonnement serait analogue pour le français *juge*, l'espagnol *juego*, etc.). Elle est nom ou verbe, et on peut admettre qu'on peut vouloir lui assigner des significations différentes selon qu'elle est nom ou verbe. Or dans le système de UG cela ne peut pas être obtenu par assignation catégorielle ambiguë, bien que celle-ci reste possible sur le plan formel dans un « langage non ambigu ».

À noter que les conséquences de l'Observation généralisée précédente sont suffisamment dévastatrices même si toutes les ambiguïtés d'assignation catégorielle n'exigent pas un corrélat différencié dans l'assignation des significations. On sait, p. ex., que pour exprimer la sémantique qu'il souhaite via LI, Montague ne distingue pas les assignations de signification des noms communs et de verbes intransitifs, de telle manière que pour *they judge* et *the judge*, l'ambiguïté d'assignation catégorielle ne conduit pas à des résultats inadéquats. Mais il suffit d'élargir les exemples à *they judge Peter* pour retrouver le problème. Si l'on souhaite exprimer sur ces exemples les interprétations qui sont censées leur être données via LI, la suite élémentaire *judge* ne peut pas être assignée de manière ambiguë à la catégorie des verbes transitifs et des verbes intransitifs.

4 Syntaxe de la langue naturelle et formules de la logique intensionnelle dans PTQ

Dans la Section 3 précédente, nous avons examiné UG en fonction des deux interrogations qui sont le fil conducteur de ce travail, à savoir l'interrogation sur l'empiricité et l'interrogation sur la compositionnalité. En ce qui concerne la première, nous avons observé que l'Idée 1 de Montague, selon laquelle il n'y aurait pas de différences théoriques

fondamentales entre langages naturels et langages artificiels des logiciens, pouvait mais ne devait pas forcément être interprétée comme une hypothèse empirique. Mais, si on l'interprétait comme hypothèse empirique, la compositionnalité devait jouer un rôle essentiel dans l'argumentation. Or, la compositionnalité dite frégiennne, en tant que telle, n'est pas centrale dans UG. Ce qui est central dans UG c'est l'homomorphisme- u ; c'est lui qui, si les quatre Remarques d'UG étaient prouvées, doit conduire à prouver l'interprétation induite. Or pour l'empiricité, le problème crucial est qu'aucune restriction n'est proposée dans UG sur l'algèbre qui doive permettre de décrire la syntaxe de la langue naturelle, et cette absence rend absente la possibilité d'interpréter empiriquement l'Idée fondamentale de Montague.

Dans l'analyse de UG, nous avons observé un point qui va revenir par la suite et qui est crucial pour comprendre un peu ce qu'est une langue naturelle. Parmi d'autres caractéristiques observationnelles généralisables, on découvre dans les langues naturelles, comme constante, l'ambiguïté d'assignation catégorielle des suites élémentaires, et ceci quel que soit le système descriptif utilisé.

Ceci est très énervant pour celui ou celle qui veut aboutir à définir un système catégoriel définissant une partition des suites élémentaires à classer dans les catégories. Et pour aboutir à ce système de classement qu'il pense idéal et qu'il a dans sa tête, pour ne pas assigner une suite – objet graphique – à plus d'une catégorie, il a recours à l'ellipse. Ainsi, dans *Donne-moi la carrée*, on dira que *carrée* est toujours un adjectif auquel on suppose accolé un nom elliptique. C'est encore une manifestation de la mise au pas de la langue naturelle par ceux qui en parlent, dans la mesure où strictement rien n'est dit sur les limites à fixer à l'ellipse et/ou sur les conditions de la récupérabilité des éléments censés être elliptiques⁹.

Nous avons remarqué dans la Section précédente que les relations entre algèbres permettant d'aboutir dans UG à l'interprétation induite sur l'univers de dénotation – et ce sont elles qui permettent de se dispenser de la LI – exigent que les suites élémentaires de la syntaxe soient assignées de manière non ambiguë aux catégories, si l'on veut obtenir des descriptions adéquates dans LI. C'est en fait l'une des rares exigences qui découlent du cadre formel, plutôt vide par rapport à la syntaxe, de UG. Mais comme il a été remarqué, le fait que l'assignation ambiguë ne soit pas possible étant donné l'adéquation exigée, ne veut pas dire que dans le cadre formel très général de UG il ne soit pas possible de traiter la question de l'assignation catégorielle ambiguë des langues naturelles, et ce pour la raison générale déjà indiquée : l'absence de contraintes sur l'algèbre qui va être utilisée pour décrire la langue naturelle.

Nous avons remarqué aussi dans la Section précédente le décalage entre UG et PTQ, alors que c'est ce dernier travail qui est la référence systématique lorsqu'on parle de la Grammaire de Montague.

⁹ Sur l'utilisation de l'ellipse pour sauver l'axiome de la monocatégorisation, cf. [Auroux 93 p. 185-202].

Cette Section sera centrée sur l'analyse de PTQ, considéré comme une concrétisation, au moins partielle, de UG. L'analyse sera conduite à partir des mêmes interrogations sur l'empiricité et la compositionnalité. Elle portera donc essentiellement sur la syntaxe de la langue naturelle et sur sa traduction en LI.

4.1 Le langage formel (pSLF)

Dans la Section 3.2.1 nous avons présenté le quintuplet qui définit un « langage », à savoir :

$$\langle A, F_\gamma, X_\delta, S, \delta_0 \rangle_{\gamma \in \Gamma, \delta \in \Delta}$$

Nous savons (cf. Section 1.1) que dans PTQ Montague ne spécifie pas un « langage non ambigu » et qu'il n'y a pas dans PTQ à proprement parler une Base de traduction, mais des éléments d'une Base de traduction. En revanche, PTQ introduit une définition récursive des indices catégoriels en syntaxe et une fonction de traduction de ces indices sur les types de LI, absentes de UG.

Le langage formel de PTQ n'est pas formel, au sens plus rigoureux donné à ce terme dans la Section 1.1, et il n'est pas non ambigu. Ce que PTQ offre est une *S~FLA* (Spécification non formelle d'un langage ambigu). Comme substitut de ce sigle nous utilisons ici *pSLF* pour *pseudo Spécification Langage Formel*, pour dénoter tout ce qui concerne en PTQ la description de la syntaxe. La justification du *p* de *pseudo* sera donnée dans le courant de l'analyse et récapitulée dans la Section 4.1.5.

On a trouvé peu de formulations claires sur le pouvoir expressif de pSLF, même s'il y en a qui orientent clairement sur le préfixe *p*.

D'après [Dowty & al. 81 p. 179], la syntaxe du fragment anglais de PTQ devrait permettre de construire les expressions non élémentaires « from the inside out », c'est-à-dire « by means of recursive definitions specifying how complex phrases are to be formed out of simpler ones. This procedure makes for a « Fregean » semantic treatment of English » [...] the syntactic rules [de pSLF] will not only concatenate input expressions, possibly with other symbols added as well, but may also perform transformation-like manipulations of the inputs. In fact, this syntax can be fairly easy converted into an equivalent grammar with a phrase structure and a simple transformational component ».

Comme on n'a jamais connu le pouvoir expressif formel d'un « transformational component »¹⁰, on a beaucoup de mal à imaginer ce qu'est le pouvoir expressif de SLP', notant par là l'objet qui est le résultat de la conversion de pSLF dans une grammaire avec une composante transformationnelle.

¹⁰ D'après [Peters & Ritchie 73] il semblerait que ce pouvoir est irrestict.

On n'essayera pas ici de mathématiser la théorie générale des langages de *Universal Grammar*. L'objectif des pages suivantes de cette Section 4.1 est beaucoup plus modeste, mais, semble-t-il, suffisamment éclairant. Il s'agit d'observer le pSLF, tel qu'il est, de PTQ et d'analyser les entités, règles et opérations qui y sont utilisées, en essayant de les caractériser en fonction de :

- la hiérarchie chomskyenne des règles et de grammaires
- les grammaires catégorielles

La démarche sera donc la suivante: une spécification X des suites-pSLF, obtenues par la machinerie X' de pSLF, peut ou doit, selon les cas, être obtenue par une règle de tel ou tel niveau dans la hiérarchie chomskyenne et/ou par tel ou tel type de grammaire catégorielle. Dans les quatre Sections suivantes on présente successivement : (i) le rappel des notions utilisées dans les sections suivantes (Section 4.1.1), (ii) la spécification des couples <règle-pSLF, opération-pSLF> de PTQ (Section 4.1.2), (iii) le pouvoir expressif de chaque couple utilisé dans PTQ (Section 4.1.3), (iv) le Vocabulaire utilisé dans pSLF (Section 4.1.4). Pour clôturer la Section 4.1, on présente dans la Section 4.1.5 tous les éléments qui justifient l'utilisation du préfixe *p* dans pSLF.

4.1.1 Notions utilisées

Étant donné un Vocabulaire de symboles, la notion générale de *règle de production* qu'on utilise par la suite peut se caractériser par le schéma suivant :

<er> --> <eg> <ed>

où :

<er> [expression résultante], <eg> [expression gauche], <ed> [expression droite]

sont des suites de symboles du Vocabulaire, et <eg> précède immédiatement <ed>.

Chaque règle doit se définir par le type de symboles que l'on peut utiliser à gauche et à droite de la flèche. On réserve *op-pSLF* et *règle-pSLF* pour désigner les opération et les règles de PTQ.

Une *règle d'interprétation* est conceptuellement le même objet que le précédent, mais conçu d'un point de vue inversé : l'entrée d'une règle de production devient sortie et vice versa. On notera donc :

<eg> <ed> <--> <er>

On sait que dans PTQ les *catégories* (plus exactement, ce sont des *indices catégoriels*) sont définies récursivement à partir des éléments basiques *e* et *t*, et de l'utilisation du slash et du double slash. On ne revient pas sur ces définitions suffisamment bien explicitées ailleurs (cf. [Dowty & al. 81 p. 182], [Chambreuil & Pariente 90 p. 70]). Seulement un sous-ensemble assez restreint des catégories définissables récursivement sont utilisées pour

caractériser les expressions élémentaires du Vocabulaire; ce sont les catégories notées habituellement *IV*, *CN*, *T*, *TV*, *IAV*, *t/t*, *IAV/T*, *IV/t*, *IV//IV*. Dans la discussion qui suit on utilisera parfois les notations en abrégées – sans slash – parfois les notations avec les mêmes catégories notées avec slash. Ainsi, indistinctement, on utilisera soit *T* soit *t/IV*. En plus des conventions habituelles, on utilisera les conventions notationnelles suivantes :

$IAV = IV/IV = ADV1$

$t/t = ADV2$

$IAV/T = PREP$

$IV/T = VEC$ (Verbe avec Enchâssées avec Complémenteur)

$IV//IV = VEI$ (Verbe avec Enchâssées à l'Infinitif)

Dans les *grammaires catégorielles* on n'a pas de règles qui vont donner le résultat de la concaténation, mais des *schémas de règles*, notés avec des variables qui vont être instanciées par les différentes catégories assignées aux expressions élémentaires et aux expressions spécifiées par application des schémas de règles. Le schéma de règles, qui sont des règles d'interprétation, pour noter l'application fonctionnelle peut être noté ainsi :

$X/Y\alpha \quad Y\beta \quad \leftarrow X\alpha\beta$

Ce schéma peut se lire : si une expression α associée à une catégorie qui instancie X/Y est placée immédiatement à gauche d'une expression β associée à une catégorie qui instancie Y , la concaténation $\alpha\beta$ sera une expression qui aura comme catégorie l'instanciation de X . L'énorme avantage de cette manière de voir les choses est que chaque expression, élémentaire ou résultante d'une concaténation, porte en elle même les avatars possibles de sa combinatoire dans les suites bien formées. Si l'on utilise une grammaire catégorielle avec l'application fonctionnelle, on doit donc exclusivement se limiter à spécifier les expressions élémentaires et les catégories auxquelles elles sont associées; on suppose le schéma de règle de l'application fonctionnelle et plus rien ne doit être spécifié : il n'y a pas de règles ad hoc qui vont dire quelle est la catégorie qui doit être associée au résultat de chaque concaténation. Il est à remarquer que, dans les vraies grammaires catégorielles utilisant l'application fonctionnelle, tout ce qui concerne X/Y , Y et la résultante X est spécifié dans ces symboles eux-mêmes, qu'ils soient des symboles monadiques ou des ensembles de traits, avec ou sans variables, avec ou sans unification : il n'y a donc pas des commentaires en LN pour expliquer ce que l'on doit faire. Il est à remarquer aussi qu'aucune opération n'est faite sur les expressions élémentaires associées à X/Y , à Y et à la résultante X : l'expression associée à X est strictement le résultat de concaténer l'expression associée à X/Y avec l'expression associée à Y .

Les règles utilisables dans une grammaire s'organisent dans une hiérarchie. On suppose qu'une grammaire détermine une dérivation ou ensemble fini de lignes, dont chaque ligne est obtenue de la précédente par application d'une règle. Si l'on se rapporte au schéma général de règles présenté ci-dessus, $\langle er \rangle$ dans la ligne n , est remplacé par $\langle eg \rangle$ $\langle ed \rangle$ dans la ligne $n+1$. Selon les contraintes sur l'entrée et la sortie, on place les règles dans une hiérarchie des contraintes de quatre niveaux (0 à 3), les règles les moins contraintes étant

placées dans le niveau 0 et les plus contraintes dans le niveau 3. Plus particulièrement et pour ce qui nous intéresse, on a :

r-0 : règle de niveau 0, qui est un niveau quasiment sans contrainte : pourvu que l'on utilise les symboles déclarés dans le Vocabulaire de la grammaire ou *nihil*, on peut tout faire sur eux, y compris les supprimer. Ce sont des règles qui spécifient des langages non récursivement décidables. Ceci veut dire qu'étant donné une suite quelconque de symboles terminaux du Vocabulaire d'une grammaire qui a incorporé une r-0, il n'existe pas d'algorithme permettant de dire si cette suite appartient ou non au langage spécifié par cette grammaire. Ce sont les seules règles de la hiérarchie qui permettent l'élimination des symboles d'une ligne de la dérivation sans qu'ils soient remplacés par d'autres symboles dans la ligne suivante.

r-1 : règle de niveau 1, niveau où déjà on ne peut plus éliminer des symboles, mais à la grande différence du niveau suivant, on peut opérer en entrée sur deux symboles adjacents (terminaux ou non terminaux). Le symbole *<er>* peut donc être *<er1><er2>*

r-2 : règle de niveau 1, niveau où *<er>* ne peut être qu'un symbole unique.

4.1.2 Spécification des couples < règle-pSLF, opération-pSLF > de PTQ

PTQ utilise des opérations-pSLF qui sont associées à des règles-pSLF. On peut répertorier seize opérations-pSLF, notées F₀ à F₁₅. Chaque opération-pSLF est associée à une ou plusieurs règles-pSLF, au nombre de 24, répertoriées V₀ à V₂₃. P. ex., F₀ est associée à seulement V₀, tandis que F₆ est associée à V₆, V₇, V₈ et V₉. Les règles-pSLF associées aux opérations-pSLF doivent stipuler, pour chaque opération-pSLF, les catégories associées aux expressions d'entrée et la catégorie associée à l'expression de sortie.

On peut énumérer les couples

< règle-pSLF, opération-pSLF >

que l'on trouve dans PTQ (désormais *couples-pSLF*¹¹) et caractériser le pouvoir expressif de chacun, en le rapportant soit à l'application fonctionnelle des grammaires catégorielles soit à une règle de la hiérarchie chomskyenne. Ceci doit permettre de caractériser le pouvoir expressif de chaque couple. Mais au-delà de ce pouvoir expressif codé dans la notation du couple < règle-pSLF, opération-pSLF >, PTQ incorpore des descriptions en langue naturelle de manipulations complexes effectuées sur les suites d'expressions terminales. Par exemple, pour F₄, on a (on reprend la formulation de [Chambreuil & Pariente 90 p. 75] :

F₄ (α, β) = (α, β')

où β' est obtenu à partir de β en remplaçant dans β la première occurrence d'une expression élémentaire d'une des catégories suivantes IV, TV, IV/T, IV//IV, par une occurrence de la troisième personne du singulier présent de cet élément.

¹¹ Caveat terminologique: un couple-pSLF (i) ≠ CBF-pSLF (ii); (i) associe une règle-pSLF – qui est elle-même un CBF-pSLF – à l'opération constitutive de la règle-pSLF.

Ces manipulations ne sont pas formalisées : ce sont des suggestions algorithmiques de recherche d'éléments dans une liste et de leur substitution par d'autres éléments. Ces manipulations sont incorporées à F_2 , F_{3-n} , F_4 , F_5 , F_{10-n} , F_{11} à F_{15} . On les notera de la manière suivante :

- si X doit être instancié par une catégorie associée à l'expression α' selon V_m dans le couple $\langle V_m, F_n \rangle$, on utilisera X^n pour indiquer que l'expression α' a été obtenue à partir de l'expression α par une manipulation incorporée à F_n .

Ainsi, p. ex., t^{11} notera la catégorie t associée à l'expression α' obtenue par le couple $\langle V_{19}, F_{11} \rangle$ plus une manipulation incorporée à la spécification de F_{11} .

4.1.3 Le pouvoir expressif des couples-pSLF

Dans le tableau ci-dessous on va caractériser le pouvoir expressif de chaque couple-pSLF utilisé dans PTQ, l'idée étant de le reformuler moyennant une règle de production de la hiérarchie chomskyenne ou une règle d'interprétation instanciant le schéma de règle de l'application fonctionnelle, le couple-pSLF et sa reformulation étant censés produire les mêmes effets. Le pouvoir expressif d'un couple-pSLF sera ainsi caractérisé par le type de règle qu'on devrait utiliser pour exprimer les mêmes résultats.

Dans le tableau ci-dessous on utilisera les notations suivantes :

$r\langle X \rangle$ note une règle – au sens général – qui possède les caractéristiques de l'exposant qui va instancier $\langle X \rangle$. Cet exposant va être 1, 2, $2c$. On a ainsi les notations suivantes :

r^1 note une règle $r-1$ dans la hiérarchie chomskyenne

r^2 note une règle $r-2$ dans la hiérarchie chomskyenne

r^{2c} note une règle $r-2$ dans la hiérarchie chomskyenne, laquelle, à la différence des r^2 peut aussi, avec les seules catégories utilisées dans PTQ auxquelles ont été attribuées des expressions élémentaires, être exprimée de manière équivalente par l'application fonctionnelle des grammaires catégorielles.

Lorsque les objets analysés ont les apparences des règles mais qu'ils ne le sont pas du fait de l'incorporation des manipulations, on utilisera le préfixe p .

$pR\langle X \rangle$ note un ensemble infini d'objets qui ont les apparences des règles, mais qui incorporent des manipulations, d'où l'utilisation du préfixe p . Les objets pR se caractérisent formellement par l'utilisation d'un couple-pSLF, qui, lui-même, aura un pouvoir expressif déterminé. PTQ utilise deux pR , qui s'appliquent n fois, avec $n \geq 0$, ces pR utilisant à leur tour un des deux couples suivants :

i $\langle V_3, n ; F_{3.n} \rangle$

ii $\langle V_{m.n}; F_{10.n} \rangle$ avec $m \in \{16, 17, 18\}$

À la grande différence des schémas des règles de la grammaire catégorielle, chaque valeur de n va introduire des modifications spécifiques. Un pR n'est donc pas une manière commode de noter plusieurs règles différentes avec une seule formule, mais bien une notation d'un ensemble infini dénombrable de règles. Un pR n'est donc pas un schéma de règles au sens où cette notion est utilisée dans les grammaires catégorielles.

Si l'on fait abstraction de l'infinitude d'opérations, on peut caractériser les couples (i) et (ii) des pR comme n'importe quel autre couple-pSLF. La caractérisation sera indiquée comme instanciation de l'exposant $\langle X \rangle$ de $pR\langle X \rangle$. On aura respectivement pour (i) et (ii) :

$$\begin{array}{l} pR\text{--}\underline{2} \\ pR\text{--}\underline{0} \end{array}$$

le soulignement de l'exposant indiquant qu'on ne caractérise pas l'ensemble infini d'opérations mais une seule parmi celles-ci.

Si PTQ utilise 16 opérations-pSLF, du fait que certaines sont associées à plusieurs règles-pSLF, on peut énumérer 24 couples-pSLF. Chacun de ces couples-pSLF est présenté dans le tableau qui suit. De gauche à droite, on trouve les colonnes suivantes :

- règle-pSLF
- opération-pSLF
- catégorie de $\langle eg \rangle$ s'il y en a une ou expression
- catégorie de $\langle ed \rangle$
- catégorie de $\langle er \rangle$

Les catégories des colonnes 3 à 5 sont directement extraites de la règle-pSLF dans la première colonne de la même ligne. Dans la sixième colonne du tableau on trouvera indiqué le pouvoir expressif du couple-pSLF déterminé par les valeurs des deux premières colonnes de la même ligne. Dans cette colonne on utilisera ainsi, selon ce qui a été dit lors de la présentation des conventions de notation, les notations suivantes :

$$r^1 \quad r^2 \quad r^{2c} \quad pr^{2c} \quad pR\text{--}\underline{2} \quad pR\text{--}\underline{0}$$

Dans la septième colonne on reformule le couple-pSLF sous la forme d'une règle de production de la hiérarchie chomskyenne et, lorsque ceci est possible avec utilisation des catégories de PTQ auxquelles ont été assignées les expressions élémentaires de PTQ, on reformule aussi le couple-pSLF par une règle d'interprétation qui instancie le schéma des règles de l'application fonctionnelle.

Chaque fois que le préfixe p est utilisé dans la sixième colonne, on utilise dans la cinquième et dans la septième, la convention X^n introduite dans la Section 4.1.2 pour noter les catégories associées à des expressions terminales résultantes de manipulations non définies formellement, Θ note l'élément vide et la suppression d'un élément.

r-pSLF	op-pSLF	eg	ed	er	type	reformulation	
V ₀	F ₀	every	CN	T	r ^{2c}	• T/CN • T ->	CN <- T every CN
V ₁	F ₁	the	CN	T	r ^{2c}	• T/CN • T ->	CN <- T the CN
V ₂	F ₂	a(n)	CN	T ²	pr ^{2c}	• T/CN • T ² ->	CN <- T ² every CN
V _{3.n}	F _{3.n}	CN	t	CN ³	pR ^{r-2}	• CN ³ ->	such that t
V ₄	F ₄	t/IV	IV	t ⁴	pr ^{2c}	• t/IV • t ⁴ ->	IV <- t ⁴ T IV
V ₅	F ₅	IV/T	T	IV ⁵	pr ^{2c}	• IV/T • IV ⁵ ->	T <- IV ⁵ TV T
V ₆	F ₆	IAV/T	T	IAV	r ^{2c}	• IAV/T • IAV	T <- IAV -> PREP T
V ₇	F ₆	IV/t	t	IV	r ^{2c}	• IV/t • IV	t <- IV -> VEC T
V ₈	F ₆	IV//IV	IV	IV	r ^{2c}	• IV//IV • IV	IV <- IV -> VEI T
V ₉	F ₆	t/t	t	t	r ^{2c}	• t/t • t	t <- t -> ADV2 t
V ₁₀	F ₇	IAV	IV	IV	r ¹	• ADV1 IV ->	IV ADV1
V ₁₁	F ₈	t	t	t	r ²	• t ->	t and t
V ₁₂	F ₉	t	t	t	r ²	• t ->	t or t

V13	F8	IV	IV	IV	r^2	\bullet IV \rightarrow	IV and IV
V14	F8	IV	IV	IV	r^2	\bullet IV \rightarrow	IV or IV
V15	F9	T	T	T	r^2	\bullet T \rightarrow	T or T
V16.n	F10.n	T	t	t^{10}	pr^{1-0}	\bullet T t \rightarrow	Θ t^{10}
V17.n	F10.n	T	CN	CN^{10}	pr^{1-0}	\bullet T CN \rightarrow	Θ CN^{10}
V18.n	F10.n	T	IV	IV^{10}	pr^{1-0}	\bullet T IV \rightarrow	Θ IV^{10}
V19	F11	t/IV	IV	t^{11}	pr^{2c}	\bullet t/IV	IV \leftarrow t^{11}
V20	F12	t/IV	IV	t^{12}	pr^{2c}	\bullet t/IV \bullet t^{12}	IV \leftarrow t^{12} \rightarrow T IV
V21	F13	t/IV	IV	t^{13}	pr^{2c}	\bullet t/IV \bullet t^{13}	IV \leftarrow t^{13} \rightarrow T IV
V22	F14	t/IV	IV	t^{14}	pr^{2c}	\bullet t/IV \bullet t^{14}	IV \leftarrow t^{14} \rightarrow T IV
V23	F15	t/IV	IV	t^{15}	pr^{2c}	\bullet t/IV \bullet t^{15}	IV \leftarrow t^{15} \rightarrow T IV

4.1.4 Le Vocabulaire de pSLF

On définit $V\text{-pSLF}$ comme étant le Vocabulaire des expressions élémentaires utilisées dans PTQ; elles ne sont donc le résultat d'aucune opération. Tout comme les couples-pSLF déjà analysés dans la Section précédente, $V\text{-pSLF}$ n'est pas homogène. Pour pouvoir caractériser sur des points importants la relation entre langage-LN et langage-pSLF on va d'abord :

- spécifier $V_t\text{-LN}^{PTQ}$, c'est-à-dire le Vocabulaire terminal dans les suites-LN qui peuvent être associées aux suites-pSLF ;
- spécifier l'organisation interne de $V\text{-pSLF}$;
- caractériser la relation entre $V\text{-pSLF}$ et $V_t\text{-LN}^{PTQ}$.

Sur cette base il devrait être possible de s'approcher d'une première caractérisation des relations entre suites-LN et suites-pSLF, point important sur lequel on reviendra dans la Section 6.

On n'introduira pas dans $V_t\text{-LN}^{PTQ}$ toutes les expressions terminales qui sont des mots graphiques dans les suites-LN. Ainsi dans $V_t\text{-LN}^{PTQ}$ on n'aura ni *up*, ni *boys*, ni plus généralement les pluriels, ni *judged* or *been* ni plus généralement les participes passés, ni plus généralement encore un mot graphique qui ne soit pas utilisé dans $V\text{-pSLF}$. On aura en revanche tous et seulement les mots graphiques qui sont soit des expressions élémentaires de PTQ (c'est le cas de p. ex. *boy* et de *run*) soit introduits par une règle-pSLF (p. ex. *each*, *to*, *a*, *an*).

$V\text{-pSLF}$ est hétérogène parce qu'il est constitué de plusieurs sous-ensembles d'expressions considérées comme élémentaires dans PTQ et qui sont introduites et manipulées de manière différente. Ces sous-ensembles sont présentés par la suite.

$Ve\text{-pSLF}$ (Vocabulaire élémentaire), constitué des expressions élémentaires déclarées en tant que telles dans ce qu'on peut nommer le Lexique-PTQ, chacune de ces expressions étant associées à une et une seule catégorie de PTQ. $Ve\text{-pSLF}$ se partitionne en trois sous-ensembles, qui suivent.

$Ve1\text{-pSLF}$, avec les expressions de $Ve\text{-pSLF}$ qui peuvent être associées bijectivement à un sous-ensemble de $V_t\text{-LN}^{PTQ}$. Les éléments de $Ve1\text{-pSLF}$ sont des mots simples de l'anglais écrit.

$Ve2\text{-pSLF}$, avec des expressions considérées élémentaires en PTQ et qui correspondent à des suites de longueur supérieure à 1 formées par des éléments de $V_t\text{-LN}^{PTQ}$. C'est le cas de, p. ex., les verbes avec une enchâssée à l'infinitif. P. ex., l'expression terminale que l'on a dans $V\text{-pSLF}$ est l'objet que nous notons *try-to*. Dans une suite-pSLF on aura ainsi, avec '+' notant la concaténation, ... + *try-to* + ... et non ... + *try* + *to* + ...

$Ve3\text{-pSLF}$, ensemble infini dénombrable des expressions $he_n, him_n, she_n, \dots$ avec $n \geq 0$.

Par ailleurs dans $V\text{-pSLF}$ on a aussi l'inclusion propre de $Vsc\text{-pSLF}$ (Vocabulaire syncatégorématique), vocabulaire qui groupe un sous-ensemble fini d'expressions considérées élémentaires introduites non dans le Lexique-PTQ, mais par une opération-pSLF. Là aussi on doit distinguer deux sous-ensembles : $Vsc1\text{-pSLF}$ et $Vsc2\text{-pSLF}$. Le premier groupe des expressions élémentaires qui sont des mots graphiques simples appartenant à $V_t\text{-LN}^{PTQ}$ et le deuxième des suites de mots graphiques.

Le tableau qui suit représente l'organisation interne de $V\text{-pSLF}$. L'indentation note l'inclusion. On donne pour chacun des cinq ensembles des exemples extraits de PTQ.

V-pSLF	
Ve-pSLF	
Ve1-pSLF	(boy, Bill)
Ve2-pSLF	(try-to)
Ve3-pSLF	(he _n)
Vsc-pSLF	
Vsc1-pSLF	(every, and)
Vsc2-pSLF	(such-that)

Autrement dit :

$$V\text{-pSLF} = Ve1\text{-pSLF} \cup Ve2\text{-pSLF} \cup Ve3\text{-pSLF} \cup Vsc1\text{-pSLF} \cup Vsc2\text{-pSLF}$$

Il est clair qu'il ne peut pas avoir de fonction bijective entre $V_t\text{-LN}^{PTQ}$ et $V\text{-pSLF}$. Cela vient de l'inclusion dans $V\text{-pSLF}$ de $Ve3\text{-pSLF}$, de $Ve2\text{-pSLF}$ et de $Vsc2\text{-pSLF}$. Par ailleurs, du fait de $Ve3\text{-pSLF}$, il n'existe pas non plus de fonction bijective entre suites-pSLF et suites bien formées appartenant à $(V_t\text{-LN}^{PTQ})^+$, c'est-à-dire à l'assemblage sans symbole vide que l'on peut former à partir de $V_t\text{-LN}^{PTQ}$. Comme il a été dit ci-dessus, on reviendra sur ces points – cruciaux pour l'évaluation de PTQ – dans la Section 6.

4.1.5 Permissivités et restrictions de pSLF

On récapitule ici les permissivités détectées dans l'analyse de pSLF :

- Absence de caractérisation de la relation R (cf.[Chambreuil & Pariente 90 p. 106-107]).
- Utilisation d'un ensemble infini d'expressions élémentaires.
- Ensemble non homogène d'expressions élémentaires, partitionnable dans des sous-ensembles avec des caractéristiques différentes selon leur traitement formel et selon leur rapport avec les expressions élémentaires de la langue naturelle.

- Définition d'un ensemble infini de catégories sans aucune restriction sur celles que l'on a le droit de s'attendre à être trouvées dans les langues naturelles pour être effectivement utilisées dans les règles.
- Possibilité d'utilisation de la multiplication des slashes en vue de différencier syntaxiquement des catégories auxquelles on veut associer un même comportement dans leur traduction en LI.
- Incorporation aux opérations de manipulations non définies formellement.
- Absence des restrictions sur les opérations-pSLF autres de celles qui viennent des manipulations utilisées et, notamment, l'utilisation d'opérations d'effacement et de permutation.

Avec les deux seuls points précédents on peut conclure à l'absence des restrictions sur les couples-pSLF, un couple-pSLF incorporant une opération-pSLF. Mais l'absence des restrictions sur les couples-pSLF vient aussi du point suivant.

- Absence des restrictions sur les règles-pSLF en relation à l'application fonctionnelle des grammaires catégorielles. Si l'on observe le tableau de la Section 4.1.3 on peut grouper les règles-pSLF en trois classes :
 - celles qui utilisent l'application fonctionnelle (p. ex. $\langle V9, F6 \rangle$, qui introduit les adverbes de phrase); on dira ici qu'elles sont des *règles-pSLF catégorielles* ;
 - celles qui n'utilisent pas l'application fonctionnelle car elles ne peuvent pas le faire soit parce que les deux expressions concaténées n'ont pas, toutes les deux, des catégories assignées (c'est p. ex. le cas des couples-pSLF qui donnent les T avec un déterminant – les trois premiers couples du tableau) soit parce que les opérations introduisent de manière syncatégorématique des expressions terminales (ce sont p. ex. les couples-pSLF qui traitent la coordination), soit parce que leur pouvoir expressif dépasse, et de beaucoup, le pouvoir expressif de l'application fonctionnelle (ce sont p. ex. les couples-pSLF qui permutent ou qui effacent); on dira ici qu'elles sont des *règles-pSLF a-catégorielles* ;
 - celles qui n'utilisent pas l'application fonctionnelle même si les expressions concaténées ont des catégories sur lesquelles en principe et formellement l'application fonctionnelle pourrait s'appliquer, mais elle ne s'applique pas dans le couple-pSLF, car le résultat ne serait pas celui qu'on cherche (c'est p. ex. le cas de $\langle V3n, F3.N \rangle$); on dira ici qu'elles sont des *règles-pSLF anti-catégorielles*.

On trouve cependant deux restrictions; elles sont dans PTQ et non dans la théorie générale des langages – inexistante, cf. Section 6 – dont PTQ devrait relever. Elle sont les suivantes :

- Aucune expression élémentaire n'est assignée à deux catégories différentes; cette restriction, autant qu'on sache, n'est formulée nulle part, mais elle découle de ce que l'on veut obtenir par la traduction des suites-pSLF en form-LI au moyen de la restriction sur

TR2 (cf. ci-dessous Section 4.2) : il s'agit d'éviter qu'une même expression élémentaire de pSLF soit traduite par deux constantes différentes en LI (cf. dans Section 3.2.1 Ob. 3).

- Aucune règle-pSLF anti-catégorielle n'est ce qu'on pourrait caractériser comme *contrairement anti-catégorielle* en entendant par là que si les deux catégories pouvaient, par application fonctionnelle donner un certain résultat, l'opération *contrairement anti-catégorielle* devrait donner un résultat autre que celui attendu par application fonctionnelle. C'est-à-dire que si l'on a bien dans les opération-pSLF l'opération anticatégorielle du type suivant :

$X/Y \quad Z \quad \leftarrow W$, avec chacune des catégories différente des deux autres

on n'a pas dans les opérations-pSLF une opération contrairement anticatégorielle du type suivant :

$X/Y \quad Y \quad \leftarrow W$, avec $X \neq W$

On espère avoir réussi à justifier le p de pSLF. L'absence de mathématisation de $pSLF$ n'est pas un vide d'un objet bien défini par ailleurs qu'il s'agirait de combler un jour. L'objet pSLF est un objet hétéroclite dont on ne perçoit aucun principe de construction généralisable : ni grammaire lexicale, malgré l'utilisation des catégories très expressives associées à un ensemble d'expressions élémentaires – car il a recours à des opérations a-catégorielles ou anti-catégorielles –, ni grammaire avec des règles de production, grâce à l'appel, dans certains cas, à l'application fonctionnelle qui les rendrait inutiles. Objet mal défini ayant recours à des ensembles infinis et à des manipulations non formelles, il est impossible de le calculer formellement; c'est-à-dire que, étant donné une suite arbitraire quelconque de symboles terminaux conjointement avec pSLF, il n'existe pas d'algorithme permettant d'associer à cette suite l'arbre dérivationnel que pSLF devrait lui assigner.

Cela est autrement grave, car, comme nous le verrons dans la Section 4.2, la traduction en form-LI n'a pas comme domaine les suites-LN associées avec une structure quelconque obtenues par un système effectivement formel et computable, ou, au moins, computable dans des limites connues ou connaissables, mais des arbres de dérivation obtenus par pSLF. Et nous aboutissons ici à l'ultime constat des permissivités de PTQ : il n'existe aucune restriction sur les relations admises entre arbres de dérivation spécifiées par pSLF et suites de la langue naturelle qui sont censées « être associées » à ces suites dérivationnelles. Tout se passe ainsi comme si la traduction entre LN et LI se faisait :

- en se donnant un système – pSLF – ou tout est admis
- en ne posant aucune restriction sur ce qui est spécifié par pSLF et les suites de la langue naturelle.

4.2 La traduction du langage-pSLF en langage-LI

Le langage-LI est spécifié moyennant des *types*, des *expressions élémentaires*, et des *CBF-LI*, couples de bonne formation des expressions LI.

L'ensemble des types est un ensemble infini dénombrable formé récursivement à partir des types basiques *e* et *t* et de l'élément *s*.

Les expressions élémentaires sont des constantes ou des variables, les deux – constantes et variables – appartenant à des ensembles infinis dénombrables. Pour tout type, il existe l'ensemble correspondant des variables.

Les couples bf-LI définissent les conditions de bonne formation des expressions non élémentaires de LI. Les form-LI sont parmi les expressions non élémentaires bien formées de LI. À partir d'une expression – élémentaire ou non – avec son type, ou d'une paire d'expressions adjacentes, chacune avec son type, un CBF-LI va définir l'expression de sortie – parfois en y ajoutant de nouvelles expressions de manière syncatégorématique mais sans manipulations sur les expressions d'entrée –, et va assigner un type à l'expression de sortie.

La traduction des suites-pSLF en form-LI se spécifie par :

- i Une fonction – TR1 – qui associe à toute catégorie spécifiable par pSLF – c'est-à-dire à toute catégorie que l'on puisse spécifier selon pSLF, que cette catégorie soit utilisée ou non pour être associée à une expression de PTQ – un type de LI.
- ii Une fonction – TR2 – qui va associer à chaque élément d'un sous-ensemble d'expressions élémentaires en pSLF, une et une seule constante dans LI. Cette constante sera associée dans LI à un type obtenu par (i) à partir de la catégorie associée dans pSLF à l'expression élémentaire dans l'entrée TR2. Par exemple *boy* de catégorie *CN* se traduit dans la constante *boy'* de type $\langle\langle s, e \rangle, t \rangle$.
- iii Une famille de fonctions – TR3 – qui vont associer les expressions élémentaires *necessarily*, *be* et les expressions élémentaires de catégorie T, à des expressions non élémentaires du langage-LI.
- iv Une famille de fonctions – TR4 – qui vont associer à chaque couple-pSLF un couple-LI, celui-ci n'étant pas nécessairement un couple bf-LI, mais devant produire en sortie une expression bien formée de LI incluant un type déterminé.

La traduction s'effectue dans un cadre de restrictions, intuitivement difficiles à respecter, et de permissivités, qui augmentent le pouvoir expressif d'ensemble du système et qui, intuitivement, sont faciles à respecter. Restrictions et permissivités sont respectivement résumées dans les deux Sections suivantes.

4.2.1 Les restrictions sur la traduction

On peut qualifier de restrictions les caractéristiques suivantes.

- La traduction des catégories pSLF en types-LI se fait selon ce qui est spécifié par la fonction TR1, très simple et très générale, qui va donner l'image unique, sous forme de type, de chaque catégorie spécifiable en pSLF. Il n'existe cependant pas une fonction bijective entre l'ensemble de catégories spécifiables en pSLF et l'ensemble des types spécifiables en LI, les éléments d'un sous-ensemble de types n'étant l'image d'aucune catégorie (p. ex. le type $\langle e, \langle e, t \rangle \rangle$) et plus d'une catégorie de pSLF peut se traduire dans un même type.
- Chaque expression élémentaire de pSLF traduite par TR2 est associée à une et une seule constante de LI. TR2 est donc bien une fonction. Mais en plus, deux expressions élémentaires différentes du langage-pSLF ne peuvent pas être associées à une même constante. TR2 est donc une fonction bijective.
- Soit EX_C l'ensemble des expressions bien formées du langage-pSLF associées à la catégorie C , que ces expressions soient élémentaires ou obtenues par un couple-pSLF. La contrainte plus forte du système établit que toute expression $ex \in EX_C$ doit être traduite par une des fonctions de traduction TR1 à TR4 dans une expression-LI de type Ty ou traduite dans une expression-LI réductible par application des couples bf-LI à une expression de type Ty , avec $TR1(C) = Ty$. Autrement dit : il y a plusieurs manières de traduire une expression-pSLF en une expression-LI, mais quelle que soit la manière adoptée pour la traduire, toutes les expressions-pSLF associées à une catégorie C doivent en dernière instance se traduire dans des expressions-LI associées à un même type Ty , ce type Ty étant celui qui est exigé par TR1.

4.2.2 Les permissivités de la traduction

On peut qualifier de permissivités les caractéristiques suivantes.

- La traduction des expressions élémentaires-pSLF en expressions-LI s'effectue sans qu'il soit possible d'établir des généralisations en fonction des différentes sous-classes des expressions élémentaires. C'est ainsi que :
 - toute expression élémentaire de V-psLF (cf. Section 4.1.4) n'est pas traduite dans une expression – élémentaire ou non – de LI. Les éléments du sous-ensemble $Vsc2$ -pSLF (p. ex. : *such-that*) ne le sont pas, de même que certains éléments du sous-ensemble $Vsc1$ -pSLF (p. ex. : *and*), alors que d'autres éléments de ce même sous-ensemble $Vsc1$ -pSLF le sont (p. ex. : *every*) ;
 - toute expression élémentaire de Ve -pSLF n'est pas associée à une expression élémentaire de LI (ex : *necessarily* et *Bill* ne le sont pas).

- La traduction des couple-pSLF en couples-LI se spécifie de manière ad hoc. En effet, tout couple-LI obtenu par TR4 n'est pas tel que le type TY' assigné à l'expression-LI résultant de la concaténation de deux expressions-LI, chacune avec son type TY^1 et TY^2 , serait le type que l'on obtiendrait si on appliquait à TY^1 et à TY^2 un des CBF-pSLF. Autrement dit, les CBF-LI obtenus par TR4 sont des couples ad hoc, chacun pouvant être défini de manière distincte sur chaque CBF-pSLF. P. ex. le couple-pSLF $\langle V3.n, F3.n \rangle$, celui-là même qui concatène un nom commun (catégorie CN) avec une relative *such-that* (catégorie t) et qui donne une expression-pSLF de catégorie CN , est associé par TR4 à un CBF-LI, qui donne le type $\langle \langle s, e \rangle, t \rangle$ à la concaténation d'une expression-LI de type $\langle \langle s, e \rangle, t \rangle$ avec une expression de type t alors qu'il n'existe pas de CBF-pSLF donnant le résultat :

$$\langle \langle s, e \rangle, t \rangle (t) = \langle \langle s, e \rangle, t \rangle.$$

- Le CBF-LI sur les expressions lambda. On a remarqué déjà qu'à tout type de LI est associé un ensemble infini dénombrable des variables de ce type. Par ailleurs, le CBF-LI sur les expressions lambda établit que pour une expression

$$\text{expr } x = \lambda \langle \text{variable} \rangle Ty1 [\langle \text{corps} \rangle] Ty2$$

on aura $\langle Ty1, Ty2 \rangle$ comme type de *expr* x . Il n'y a aucune restriction sur $Ty1$ et sur $Ty2$, de telle manière qu'avec les variables à disposition, on peut fabriquer des expressions lambda qui peuvent recevoir les types les plus variés, ce qui est fort commode lorsqu'on veut obtenir un type donné pour respecter la contrainte forte présentée ci-dessus.

Notons enfin que c'est l'existence de la famille TR4 (traduction des CBF-pSLF en CBF-LI) celle qui rend possible que l'entrée des traductions soit la suite dérivationnelle associée à chaque suite terminale du langage-pSLF. On a dans chaque ligne de la dérivation, l'information enregistrée du couple-pSLF qui a permis de l'obtenir, et c'est sur cette entrée que la traduction de TR4 va opérer.

4.3 Analyse d'exemples

Par le biais d'exemples groupés en trois classes, on essayera de déterminer l'adéquation descriptive du système PTQ, en cherchant à détecter ses limites. Étant donné l'indéfinition formelle du système, il est impossible d'aboutir à une caractérisation précise de ce que PTQ peut ou ne peut pas faire par rapport à l'anglais, et ceci dans les mêmes domaines que ceux où l'anglais a été exploré dans PTQ. Mais les classes d'exemples devraient permettre d'induire la caractérisation de certaines limites de PTQ, pour pouvoir explorer ensuite la possibilité ou non de les dépasser. Les classes d'exemples sont :

- Suites-pSLF qui ne sont pas bien formées (ou dans la terminologie de [Chambreuil & Pariente 90], qui sont spécifiées par pSLF comme des suites significatives et qui ne le

sont pas) et dont la spécification erronée découle d'une mauvaise formulation des manipulations incorporées aux règles.

- Suites-pSLF auxquelles on donne une mauvaise traduction en form-LI et qui n'utilisent pas comme vocabulaire *Ve1-pSLF*.
- Suites qui devraient être spécifiées par pSLF et qui ne le sont pas, et qui utilisent comme vocabulaire *Ve1-pSLF*.

Sur la première classe d'exemples, cf. [Chambreuil & Pariente 90 §6.1] et [Chambreuil 89 p. 64-68]. En général, il s'agit des manipulations mal définies qui, dans une interprétation stricte, donnent de mauvais résultats. On ne s'attardera pas sur ces exemples car ils viennent de la partie non formelle de PTQ, sur laquelle il nous semble inutile d'épiloguer.

La deuxième classe d'exemples semble plus inquiétante pour un système dont le souci est justement d'exprimer la sémantique d'un langage qui se veut proche de la langue naturelle. Les exemples détectés (non exhaustivement) concernent :

- les anaphores dans la phrase simple ;
- le complémenteur concaténé au verbe de la matrice.

L'anaphore dans la phrase simple est réglée par $\langle V_{16.n}, F_{10.n} \rangle$ et par l'utilisation des éléments de *Ve3-pSLF*. Ces éléments (*he₀*, *him_n* ...) sont des objets non associés de manière systématique à ce qui est observable. Si l'on se donne une règle de substitution comme celle suggérée dans [Chambreuil & Pariente 90 p. 100-101], les suites suivantes

Bill loves him
He loves him

peuvent recevoir une interprétation qui en anglais ne peut correspondre qu'à respectivement

Bill loves himself
He loves himself

L'utilisation de *asserts-that*, élément de *Ve2-pSL*, exemple de complémenteur concaténé au verbe de la matrice, induit de mauvaises interprétations dans la coordination des enchâssées et pour l'emplacement de l'adverbe.

Soit les suites terminales de pSLF suivantes, qui vont illustrer la coordination des enchâssées :

Peter asserts-that John likes Suzanne and Robert asserts-that Albert likes Elisabeth
Peter asserts-that John likes Suzanne and Robert walks

Ces suites sont analysées par PTQ avec deux suites dérivationnelles différentes qui vont donner deux form-LI différentes selon que la phrase après le *and* soit coordonnée avec l'enchâssée qui précède immédiatement ou qu'elle soit coordonnée avec la principale. Ainsi, p. ex. pour la première de ces phrases, on aura, en abrégé, les deux analyses suivantes

Peter asserts ((that John likes Suzanne) and (Robert asserts that Albert likes Elisabeth))

((Peter asserts that John likes Suzanne) and (Robert asserts that Albert likes Elisabeth))

Avec la première de ces analyses et la traduction sémantique correspondante, on dit que Peter affirme deux choses, l'une étant que *John likes Suzanne* et l'autre que *Robert thinks that Albert likes Elisabeth*. Cette analyse est observationnellement fautive. Si la phrase en question est censée « correspondre » à la phrase anglaise

Peter thinks that John likes Suzanne and Robert thinks that Albert likes Elisabeth

son analyse est erronée car la phrase anglaise n'est interprétée que comme dans la deuxième analyse, dont la glose serait quelque chose comme : il y a eu deux événements, dont l'un a consisté à ce que *Peter asserts that John likes Suzanne* et l'autre à ce que *Robert asserts that Albert likes Elisabeth*. Exactement le même type de remarques s'applique à la deuxième phrase présentée ci-dessus.

La généralisation est facile à établir : la coordination des enchâssées avec complémentateur, en anglais – mais en français aussi –, exige que les phrases cordonnées soient, toutes, introduites par un même complémentateur. On a ainsi

((que P) *and* (ou *et*) (que P))

et non

((que P) *and* (ou *et*) P)

Or ceci est impossible à obtenir dans PTQ du fait de l'introduction dans *Ve-pSLF* des expressions terminales comme *asserts-that* où le complémentateur est rendu solidaire du verbe.

Soit la suite terminale de pSLF suivante, qui va illustrer l'emplacement des adverbes

Peter asserts-that voluntarily John does not love Mary

Étant donné le système pour obtenir les adverbes sur le syntagme verbal, la suite dérivationnelle pour décrire la phrase précédente suppose la ligne

voluntarily asserts-that John does not love Mary

C'est l'opération de permutation du couple <V₁₀, F₇> qui est censée placer l'adverbe dans la bonne position, c'est-à-dire après le verbe. Mais ici encore, grâce au fait que la suite terminale est *asserts-that*, *voluntarily* est située dans une position qui n'est pas syntaxiquement erronée – la suite de l'anglais « correspondante » à la suite traitée est parfaitement bien formée – mais dont l'interprétation est que *voluntarily* porte sur *does not love Mary* – et non sur *asserts-*, comme le fera porter la traduction en LI.

Nous ne nous attarderons pas non plus sur ce type de difficultés qui sont artificiellement introduites du fait même d'utiliser des concaténations factices, inexistantes dans le langage objet étudié. Notre analyse se concentre désormais sur des exemples de la troisième classe, correspondant donc à des suites qui utilisent *Vel-pSLF*, qui devraient être spécifiées par *pSLF* et qui ne le sont pas.

Bien entendu, ces suites ne seront pas cherchées dans toutes les suites de l'anglais et pour toutes les constructions aujourd'hui répertoriées de l'anglais. On considère, dans les Exemples 1 à 3 dont l'énumération suit, des suites dans strictement le même type de structures que celles traitées dans PTQ, à savoir :

- Exemple 1 Syntagmes prépositionnels comme des modificateurs des noms
- Exemple 2 Lemmes verbaux avec plus d'une sous-catégorisation
- Exemple 3 Emplacement des adverbes

Chacun de ces exemples est développé par la suite.

Exemple 1 Syntagmes prépositionnels comme modificateurs des noms

On sait qu'avec le couple $\langle V_6, F_6 \rangle$, PTQ permet de construire les syntagmes prépositionnels, et par le couple $\langle V_{10}, F_7 \rangle$ il les utilise comme modificateurs des syntagmes verbaux. Mais dans PTQ on ne peut pas spécifier la suite-*pSLF* qui devrait « correspondre » à la suite de l'anglais

the boy in the park

Il n'y a pas parmi les couples-*pSLF* un couple qui devrait permettre de permuter le syntagme prépositionnel modifieur et de passer à droite du nom, de manière à obtenir la description de la phrase précédente à partir de la ligne suivante de la dérivation

the in the park boy

Exemple 2 Lemmes verbaux avec plus d'une sous-catégorisation

Avec PTQ on peut obtenir les descriptions « correspondantes » à

Bill eats the fish

Bill runs

mais on ne peut pas obtenir la description pour

Bill eats

Bille runs a programme

La généralisation est, ici aussi, facile à établir : aucune expression n'est associée dans *Vel-pSLF* à deux catégories différentes. Et cela n'est pas un hasard conjoncturel de la

description modèle réduit de PTQ : c'est bien le résultat d'une des rares restrictions du système (cf. Section 4.1.5).

Exemple 3 Emplacement des adverbes

En anglais les adverbes dits de phrase (identifiés ci-dessous par ADV2) et les adverbes dits de syntagme verbal (identifiés ci-dessous par ADV1) peuvent s'attester dans beaucoup de positions dans la phrase, mais, si l'on reste aux deux positions choisies par PTQ – à gauche de la phrase et à droite du syntagme verbal –, les deux peuvent s'attester dans les deux positions. Ainsi on peut avoir pour les ADV1 (rappel : les syntagmes prépositionnels sont assimilés à des adverbes ADV1, cf. ci-dessus l'Ex. 1) :

In the park Peter runs slowly
Peter runs in the park slowly

et pour les ADV2

Peter runs necessarily in the park
Necessarily Peter runs in the park

Or PTQ ne spécifie des suites terminales « correspondantes » pour ADV1 qu'à droite du syntagme verbal et pour ADV2 qu'à gauche de la phrase, de telle manière que seul le deuxième membre de chacun des couples précédents peut être spécifié.

5 Bilan sur le traitement de LN dans PTQ

À partir de la présentation des exemples dans la Section 4.3, on pourrait être tenté de conclure à une sorte de radicale inadéquation descriptive de PTQ, car la présentation laisse penser que la mécanique même de PTQ, avec les rares restrictions qu'elle introduit, est intrinsèquement incapable de les traiter. Cela semblerait particulièrement évident pour traiter les lemmes verbaux catégoriellement ambigus – c'est-à-dire qui nécessitent d'être associés à deux catégories syntaxiques différentes – étant donné les restrictions de PTQ sur TR2 : une expression élémentaire de pSLF doit être traduite sur une seule constante de LI avec le type qui lui correspond selon TR1. Or, ainsi conçue, cette radicale inadéquation descriptive n'existe pas.

Après présentation des conventions de notation (Section 5.1), on va essayer de montrer (Section 5.2) que les exemples groupés en Exemple 1 à Exemple 3 de la Section précédente peuvent, tous, être traités par la machinerie de PTQ, et en respectant à la lettre ses rares restrictions; mais en utilisant son éventail de permissivités. On ne traitera pas dans cette section et dans les suivantes des manipulations introduites dans les Op-pSLF. En effet, si l'on accepte l'idée selon laquelle dans une langue LN on peut tout dire, il sera toujours possible de substituer à la description actuelle de la manipulation une autre

manipulation que l'on va juger plus adéquate; mais on observera que l'ancienne description, aussi bien que la nouvelle, ne peuvent pas être calculées.

Dans la Section 5.3 les exemples de la Section 5.2 seront étendus – toujours dans les limites de la problématique de PTQ. On essayera de dégager un cadre général où tout problème rentrant dans ce cadre général peut trouver solution étant donné les permissivités de PTQ.

5.1 Notation

Dans les présentations qui suivent jusqu'à la fin de ce document, on utilise les conventions de notation suivantes :

CS : catégories syntaxique de pSLF

Ty : type de LI

c_X : constante c de typex dans LI

va_X : variable va de typex dans LI

Étant donné X et X' , X' note la traduction en LI du X dans pSLF, X , X' prenant les valeurs suivantes :

- α : expression-pSLF (élémentaire ou non), avec sa catégorie syntaxique
- α' : expression-LI (élémentaire ou non), avec le type qui doit lui être associé selon la restriction imposée par TR1.
- α CS : expression-pSLF (élémentaire ou non), avec sa catégorie syntaxique détachée
- α' Ty : expression-LI (élémentaire ou non), avec le type détaché qui doit lui être associé selon la restriction imposée par TR1 sur CS.
- CA : catégorie de pSLF
- $\langle CA \rangle$: type qui doit être assigné selon TR1 (p. ex. si $CA = IV$, $\langle CA \rangle = \langle \langle s, e \rangle, t \rangle$).

On va travailler plus particulièrement avec les trois Types suivants, pour lesquels on établit les conventions de notation correspondantes:

Ty1 : $\langle \langle s, e \rangle, t \rangle$ [c'est-à-dire que : Ty1 = IV' = CN']

Ty2 : $\langle \langle s, \langle \langle s, e \rangle, t \rangle, t \rangle$ [c'est-à-dire que : Ty2 = T']

Ty3 : $\langle \langle s, \langle \langle s, \langle \langle s, e \rangle, t \rangle, t \rangle, \langle \langle s, e \rangle, t \rangle \rangle$ [c'est-à-dire que : Ty3 = $\langle (IV/T)' \rangle$]

Il sera ainsi possible de noter $cTy1$, $cTy2$, $cTy3$, pour noter, respectivement, des constantes de types Ty1, Ty2 et Ty3, et de manière analogue pour les variables, on notera $vaTy1$, $vaTy2$, $vaTy3$.

Comme la lecture des types est particulièrement mal commode, pour caractériser constantes et variables, on s'autorise aussi à utiliser la convention C, CA' en indice. Ainsi on a, p. ex. :

$$\begin{aligned} cTy2 &= cT' \\ vaTy1 &= va\langle IV' \rangle \end{aligned}$$

De même dans la notation de types, on pourra le faire comme ci-dessus – en utilisant les primitifs e, t, s – ou bien en utilisant la convention CA' . Ainsi :

$$Ty2 = \langle \langle s, \langle \langle s, e \rangle, t \rangle \rangle, t \rangle = \langle \langle s, IV' \rangle, t \rangle$$

Dans tous les nouveaux couples qui vont être présentés (désormais *ncouple-pSLF*), si on utilise la même opération et/ou la même règle que dans pSLF, on conserve les identifications qui sont données dans le tableau de la Section 4.1.3. Si on y introduit des variations mineures – toujours explicitées – ces variations sont notées à l'aide du suffixe b . Ainsi, p. ex., le *ncouple-pSLF* $\langle V10b, F7 \rangle$ va introduire une modification sur les catégories des expressions d'entrée en conservant la même opération : le $F7$ du couple $\langle V10b, F7 \rangle$ est exactement le même $F7$ que le $F7$ du couple $\langle V10, F7 \rangle$, alors que la règle $V10b$ n'utilise pas les mêmes catégories que la règle $V10$.

La notation d'un *ncouple-pSLF* sera

$$\alpha CS^i \quad \beta CS^j \quad \longrightarrow \quad \gamma CS^k$$

où αCS^i et βCS^j notent les expressions d'entrée et γCS^k le résultat obtenu par l'opération et la règle du *ncouple-pSLF* sur αCS^i et βCS^j .

Chaque *ncouple-pSLF* sera associé, comme il se doit selon TR4, à une traduction. Étant donné *ncouple-pSLF_i*, on va noter $Tr(ncouple-pSLF_i)$ la traduction exigée par TR4 qui va associer les expressions-pSLF aux expressions-LI.

Étant donné le *ncouple-pSLF_i* :

$$\alpha CS^i \quad \beta CS^j \quad \longrightarrow \quad \gamma CS^k$$

$Tr(ncouple-pSLF_i)$ se spécifie ainsi :

$$\alpha ' \quad \beta ' \quad \longrightarrow \quad \langle \text{[expression LI]} \rangle$$

Il s'agira de montrer pour chaque paire :

- *ncouple-pSLF_i*
- $Tr(ncouple-pSLF_i)$

que l'on a bien respecté :

- $TR1(\alpha CS^i) = \alpha '$
- $TR1(\beta CS^j) = \beta '$

et que l'[*expression LI*] a comme type TY tel que $TR1(CS^k) = TY$.

Pour montrer le respect des contraintes sur la traduction, on présentera des lignes successives où chacune s'obtient de la précédente par utilisation d'un CBF-LI et/ou d'un remplacement d'indices de constantes et variables, ce remplacement d'indices découlant des conventions de notation que l'on vient de présenter. Une ligne pleine au-dessous des symboles de la ligne précédente note la répétition de ces symboles. Ainsi les lignes n et $n+1$:

n	α	β	γ
$n+1$	<hr/>		κ

notent la même succession que les lignes n et $n'+1$ suivantes :

n	α	β	γ
$n'+1$	α	β	κ

5.2 Reconsidération des Exemples 1 à 3 de la Section 4.3

On reconsidère dans cette Section les cas groupés dans les Exemples 1 à 3 de la Section 4.3, c'est-à-dire :

- syntagmes prépositionnels comme modificateurs des noms (Section 5.2.1) ;
- lemmes verbaux avec plus d'une sous-catégorisation (Section 5.2.2) ;
- emplacement des adverbes (Section 5.2.3).

Dans la formulation des ncouples-pSLF et des traductions afférentes, on respecte strictement les restrictions de PTQ. Mais il est clair aussi que l'on jouira des mêmes permissivités de PTQ et notamment :

- traductions ad hoc de certaines expressions élémentaires de p-SLF ;
- obtention de l'expression LI en sortie d'une $Tr(ncouple\text{-}pSLF_i)$ autrement que par les couplesbf-LI. Étant donné :

$Tr(ncouple\text{-}pSLF_i)$:

$$\alpha' \quad \beta' \quad \longrightarrow \quad \langle [expression\ LI] \rangle$$

on a les mêmes exigences que dans PTQ sur α' , sur β' sur $\langle [expression\ LI] \rangle$, mais on n'obtient pas forcément $\langle [expression\ LI] \rangle$ par un CBF-LI sur α' et sur β' (cf. la traduction du CBF-pSLF $\langle V_{3.n}, F_{3.n} \rangle$), même si, comme dans toutes les traductions de PTQ, les expressions α' et β' sont incorporées dans l' $\langle [expression\ LI] \rangle$.

Par ailleurs, on jouira aussi de la définition des types, des variables et des constantes dans LI, et notamment de disposer des variables pour tous les types spécifiables dans LI.

5.2.1 Syntagmes prépositionnels comme modificateurs des noms

Un ncouple-pSLF et sa traduction sont les seules extensions nécessaires à introduire dans pSLF pour exprimer les syntagmes prépositionnels modificateurs des noms, comme p. ex. dans *boy in the park*.

$$\begin{array}{lll} \text{ncouple-pSLF}_i \langle V_{10b}, F_7 \rangle: & \alpha_{IAV} & \beta_{CN} \longrightarrow (\beta \alpha) \\ \text{Tr(ncouple-pSLF}_i): & \alpha' & \beta' \longrightarrow [\alpha' (^{\beta'})] \end{array}$$

L'expression $[\alpha' (^{\beta'})]$ a bien CN' comme type :

Étant donné

$$\alpha' \langle \langle s, Ty1 \rangle, \langle Ty1 \rangle \rangle \quad \beta' Ty1 \quad \longrightarrow [\alpha' (^{\beta'})]$$

$$\begin{aligned} \text{on a} \quad [\alpha' (^{\beta'})] &= [\alpha' \langle \langle s, Ty1 \rangle, \langle Ty1 \rangle \rangle (^{\beta' Ty1})] \\ &\quad [\text{_____} (\beta' \langle s, Ty1 \rangle)] \\ &\quad [\text{_____}] Ty1 \end{aligned}$$

et $Ty1 = CN'$

Exemple : *boy in the park*

On a en entrée :

- en pSLF (in the park)IV/IV boyCN
- en IL (in the park)'<(IV/IV)'> boy'<CN'>

Par ncouple-pSLF_i et Tr(ncouple-pSLF_i) on obtient en sortie :

- en pSLF boyCN (in the park)IV/IV
- en IL [(in the park)'<(IV/IV)'> (^boy'CN')<CN'>

C'est-à-dire que, grâce à son pouvoir expressif, le couple-pSLF arrive à permuter les expressions-pSLF, tandis que la traduction en IL est tenue à l'écart de cette opération sur les suites de la dérivation et elle maintient en IL l'expression-IL *(in the park)'* comme foncteur sur *boy'*.

5.2.2 Lemmes verbaux avec plus d'une sous-catégorisation

On considère ici le cas de *eat* qui peut être soit transitif soit intransitif. Les extensions à introduire à PTQ sont les suivantes :

- assignation de *eat* dans Ve-pSLF à IV//T et non à IV/T
- incorporation de *vac1* [*vac* mnémotechnique de *vacuum*] à Ve-pSLF de catégorie IV/(IV//T) avec une traduction ad hoc :

$$\text{Trad}(vac1) = \lambda va1 \langle s, Ty3 \rangle [\sim va1 (^{\sim va} \langle Ty2 \rangle)].$$

• définition de deux nouveaux couples avec leur traduction :

– ncouple-pSLF_j : $\text{vac1}_{IV/(IV//T)} \quad \alpha_{IV//T} \longrightarrow \alpha_{IV}$
 $\text{Tr}(\text{ncouple-pSLF}_j) : \quad \text{vac1}' \quad \alpha'_{<(IV//T)>} \longrightarrow \text{vac1}' (^{\alpha'})$

– ncouple-pSLF_k = <V_{5b}, F₅>

$\text{Tr}(\text{ncouple-pSLF}_k) = \text{Tr}(<V_5, F_5>)$

[où dans V₅ on a IV//T à la place de IV/T dans V₅].

On doit vérifier que

- i $\text{Type}(\text{TR1}(IV/(IV//T))) = \text{Type}(\lambda \text{val}_{<s, \text{Ty3}>} [\sim \text{val} (^{\text{va}_{<\text{Ty2}>}})])$
- ii $\text{Type}(\text{vac1}' (^{\alpha'})) = <IV'>$

Vérification de (i)

Étant donné $\text{TR1}(IV/(IV//T))$, on a :

$$\begin{aligned} \text{TR1}(IV/(IV//T)) &= <<s, <(IV//T)'>>, <IV'>> \\ &= <<s, \text{Ty3}>, <IV'>> \end{aligned}$$

Étant donné $\text{Tr}(\text{vac1})$ on a :

$$\begin{aligned} &\lambda \text{val}_{<s, \text{Ty3}>} [\sim \text{val} (^{\text{va}_{<\text{Ty2}>}})] \\ &\quad \text{_____} [\text{val}_{<\text{Ty3}>} (^{\text{va}_{<s, \text{Ty2}>}})] \\ &\quad \text{_____} [\text{val}_{<<s, \text{Ty2}>, <IV'>>} \text{_____}] \\ &\quad \text{_____} [\text{_____}]_{<IV'>} \\ &\quad \text{_____}]_{<<s, \text{Ty3}>, <IV'>>} \end{aligned}$$

Vérification de (ii)

Étant donné :

$$\text{vac1}' \alpha'_{<(IV//T)>} \longrightarrow \text{vac1}' (^{\alpha'})$$

on a

$$\begin{aligned} \text{iii} \quad &\lambda \text{val}_{<s, \text{Ty3}>} [\sim \text{val} (^{\text{va}_{<\text{Ty2}>}})] (^{\text{c}_{<(IV//T)'>}}) \\ &\quad \text{_____} (\text{c}_{<s, <(IV//T)'>>}) \\ &\quad \text{_____} (\text{c}_{<s, \text{Ty3}>}) \\ &\quad \quad \text{_____} [\sim \text{c}_{<s, \text{Ty3}>} (^{\text{va}_{<\text{Ty2}>}})] \\ &\quad \quad \text{_____} [\text{c}_{<\text{Ty3}>} (\text{va}_{<s, \text{Ty2}>})] \\ &\quad \quad \text{_____} [\text{c}_{<<s, \text{Ty2}>, <IV'>>} (\text{va}_{<s, \text{Ty2}>})] \\ &\quad \quad \text{_____}]_{<IV'>} \end{aligned}$$

Exemple : *[John] eats*

On a en entrée :

- en pSLF: $\text{vac1}_{IV/(IV//T)} \quad \text{eat}_{(IV//T)}$
- en IL $\text{vac1}' \quad \text{eat}'_{\langle(IV//T)'\rangle}$

Par ncouple-pSLF_j et Tr(ncouple-pSLF_j) en sortie :

- en pSLF, on obtient eat_{IV}
- en IL, en remplaçant dans (iii) c par eat , on obtient (iv).

iv $[\text{eat}'_{\langle\langle s, \text{Ty2} \rangle, \langle IV' \rangle \rangle} (\text{va}_{\langle s, \text{Ty2} \rangle})]_{\langle IV' \rangle}$

La traduction de *John eats* peut se gloser ainsi : *John mange quelque chose mais on ne sait pas quoi*, la chose en question étant exprimée par la variable $(\text{va}_{\langle s, \text{Ty2} \rangle})$.

Le ncouple-pSLF_k = $\langle V5b, F5 \rangle$ et sa traduction vont permettre de continuer à traiter *John eats the fish* comme dans les formulations actuelles.

5.2.3 Emplacement des adverbes

Deux situations sont considérées :

- l'emplacement des ADV1 à la gauche immédiate de la phrase, afin de pouvoir obtenir :

Voluntarily John losed the pen

avec la même traduction LI que

John losed the pen voluntarily

- l'emplacement des ADV2 à la gauche immédiate du syntagme verbal, afin de pouvoir obtenir :

John necessarily runs

avec la même traduction LI que

Necessarily John runs

Emplacement des ADV1

On considère ici les adverbes associés dans PTQ à la catégorie IV/IV, comme, p. ex., *voluntarily*. Les extensions à introduire à PTQ sont les suivantes :

- incorporation de *flo1* (*flo* mnémotechnique de *flottant*) à Ve-pSLF associé à la catégorie IV///IV avec une traduction ad hoc
- $\text{Trad}(\text{flo1}) = \lambda \text{va1}_{\langle s, \langle IV' \rangle \rangle} [\text{va}_{\langle (IV/IV)' \rangle} (\text{va1})]$

- définition de deux nouveaux couples avec leur traduction :

$$\begin{aligned}
 & - \text{ncouple-pSLF}_l \quad \text{flo1}_{IV///IV} \quad \alpha_{IV} \quad \rightarrow \quad \alpha_{IV} \\
 & - \text{Tr}(\text{ncouple-pSLF}_l) \quad \text{flo1}' \langle (IV///IV)' \rangle \alpha' \langle IV' \rangle \quad \rightarrow \quad \text{flo1}' (\wedge \alpha' \langle IV' \rangle) \\
 & - \text{ncouple-pSLF}_m \quad \alpha_{IV/IV} \quad \beta_t \quad \rightarrow \quad (\alpha \beta)_t \\
 & - \text{Tr}(\text{ncouple-pSLF}_m) \quad \alpha' \langle (IV/IV)' \rangle \beta'_t \quad \rightarrow \quad \lambda va \langle (IV/IV)' \rangle [\beta'_t] (\alpha' \langle (IV/IV)' \rangle)
 \end{aligned}$$

Observations :

- On conserve de PTQ toutes les expressions de catégorie IV/IV avec leur traduction et le couple-pSLF $\langle V_{10}, F_7 \rangle$ avec sa traduction.
- On doit vérifier (i), (ii), (iii) suivants, la vérification de (i) et (ii) concernant $\text{Tr}(\text{couple-pSLF}_l)$, tandis que (iii) concerne $\text{Tr}(\text{ncouple-pSLF}_m)$:
 - $\text{Type}(\text{TR1}(IV///IV)) = \text{Type}(\lambda va1_{\langle s, \langle IV' \rangle \rangle} [va \langle (IV/IV)' \rangle (va1)])$
 - $\text{Type}(\text{flo1}'(\wedge \alpha' \langle IV' \rangle)) = \langle IV' \rangle$
 - $\text{Type}(\lambda va \langle (IV/IV)' \rangle [\beta'_t] (\alpha' \langle (IV/IV)' \rangle)) = \langle t \rangle$

Vérification de (i)

Rappel : $\text{TR1}(IV///IV) = \langle \langle s, \langle IV' \rangle \rangle, \langle IV' \rangle \rangle$

Étant donné $\text{Trad}(\text{flo1})$ on a :

$$\begin{aligned}
 & \lambda va1_{\langle s, \langle IV' \rangle \rangle} [va \langle (IV/IV)' \rangle (va1)] \\
 & \underline{\hspace{10em}} [\hspace{10em}] \\
 & \hspace{15em} \langle IV' \rangle \\
 & \underline{\hspace{10em}}] \\
 & \hspace{15em} \langle \langle s, \langle IV' \rangle, \langle IV' \rangle \rangle
 \end{aligned}$$

Vérification de (ii)

Étant donné :

$$\text{flo1}' \langle (IV///IV)' \rangle \alpha' \langle IV' \rangle \quad \rightarrow \quad \text{flo1}' (\wedge \alpha' \langle IV' \rangle)$$

on a :

$$\begin{aligned}
 & \lambda va1_{\langle s, \langle IV' \rangle \rangle} [va \langle (IV/IV)' \rangle (va1)] \quad (\wedge c \langle IV' \rangle) \\
 & \underline{\hspace{10em}} [\hspace{10em}] \quad (c \langle s, \langle IV' \rangle \rangle) \\
 & \hspace{15em} [\hspace{10em} (c \langle s, \langle IV' \rangle \rangle)] \\
 & \hspace{15em} [va \langle \langle s, \langle IV' \rangle \rangle, \langle IV' \rangle \rangle (c \langle s, \langle IV' \rangle \rangle)] \\
 & \hspace{15em} [\hspace{10em}] \\
 & \hspace{15em} \langle IV' \rangle
 \end{aligned}$$

Vérification de (iii)

Étant donné :

- $\alpha' \langle (IV/IV)' \rangle > \beta'_t \rightarrow \lambda va \langle (IV/IV)' \rangle [\beta'_t] (\alpha' \langle (IV/IV)' \rangle)$
- $[\beta'_t] = [\dots va \langle (IV/IV)' \rangle \dots]_{\langle t \rangle}$ où $va \langle (IV/IV)' \rangle$ a été introduite par $Tr(ncouple-pSLF_m)$

on a :

$$\lambda va \langle (IV/IV)' \rangle [\dots va \langle (IV/IV)' \rangle \dots]_{\langle t \rangle} (c \langle (IV/IV)' \rangle) \\ [\dots c \langle (IV/IV)' \rangle \dots]_{\langle t \rangle}$$

On remarque que $ncouple-pSLF_l$ fait disparaître $flo1$ de la dérivation de l'expression finale, tandis qu'avec la variable $va \langle (IV/IV)' \rangle$ il introduit un marque-place dans la forme-LI. Grâce à $Tr(ncouple-pSLF_m)$ ce marque-place sera occupé suivant l'abstraction lambda par la constante dans laquelle l'expression ADV1 a été traduite, de telle manière que l'on obtient pour ADV1 situé à gauche de la phrase verbale la même traduction que lorsque l'expression est située à droite du syntagme verbal. On remarque enfin que le rôle du $ncouple-pSLF_m$ est de concaténer l'expression de l'ADV1 en tête de l'expression de la phrase en conservant pour le tout la catégorie t .

Emplacement des ADV2

On considère ici l'adverbe associé dans PTQ à la catégorie t/t , c'est-à-dire *necessarily*. Les extensions à introduire à PTQ sont les suivantes :

- incorporation de $flo2$ à $Ve-pSLF$ associé à la catégorie $IV//IV$ avec une traduction ad hoc
- $Trad(flo2) = \lambda va1_{\langle s, \langle IV' \rangle \rangle} [\forall va]$
- définition de deux nouveaux couples avec leur traduction :
 - $ncouple-pSLF_n \quad flo2_{IV//IV} \quad \alpha_{IV} \quad \rightarrow \quad (flo2 \alpha) IV$
 - $Tr(ncouple-pSLF_n) \quad flo2' \langle (IV/IV)' \rangle \quad \alpha' \langle IV' \rangle \quad \rightarrow \quad flo2'(\wedge \alpha' \langle IV' \rangle)$
 - $ncouple-pSLF_o \quad necessarily_{t/t} \quad \alpha_t \quad \rightarrow \quad \alpha^0$
 - $Tr(ncouple-pSLF_o) = Tr(\langle V9, F6 \rangle)$
- manipulation : α^0 s'obtient à partir de α en remplaçant $flo2$ par *necessarily*.

Observations :

- on conserve de PTQ *necessarily* $\langle t/t \rangle$ et sa traduction, le couple $(\langle V9, F6 \rangle)$ et sa traduction.
- on doit vérifier (i) et (ii) suivants :
 - i $Type(TR1(IV//IV)) = Type(\lambda va1_{\langle s, \langle IV' \rangle \rangle} [\forall va])$
 - ii $Type(flo2'(\wedge \alpha' \langle IV' \rangle)) = \langle IV' \rangle$.

Vérification de (i)

Rappel : $TR1(IV//IV) = \langle \langle s, \langle IV' \rangle \rangle, \langle IV' \rangle \rangle$

Étant donné $Trad(flo2)$ on a :

$$\frac{\lambda va_{\langle s, \langle IV' \rangle \rangle} [\forall va_{\langle s, \langle IV' \rangle \rangle}] \quad \frac{[va_{\langle IV' \rangle}]}{\langle \langle s, \langle IV' \rangle \rangle, \langle IV' \rangle \rangle}}{[va_{\langle IV' \rangle}]} \quad \langle \langle s, \langle IV' \rangle \rangle, \langle IV' \rangle \rangle$$

Vérification de (ii)

Étant donné :

$$flo2' \langle (IV/IV)' \rangle \alpha' \langle IV' \rangle \quad \rightarrow \quad flo2' (\wedge \alpha' \langle IV' \rangle)$$

on a :

$$\text{iii} \quad \frac{\lambda va_{\langle s, \langle IV' \rangle \rangle} [\forall va_{\langle s, \langle IV' \rangle \rangle}] \quad (\wedge c \langle IV' \rangle)}{[\forall c_{\langle s, \langle IV' \rangle \rangle}] \quad (c_{\langle s, IV' \rangle})} \quad (c_{\langle s, IV' \rangle})$$

Exemple : $flo2 \text{ run}$

On a en entrée :

- en pSLF : $flo2_{IV/IV} \text{ run}_{IV}$
- en IL : $flo2' \langle (IV/IV)' \rangle \text{ run}' \langle IV' \rangle$

Par $ncouple-pSLF_n$ et $Tr(ncouple-pSLF_n)$ en sortie :

- en pSLF, on obtient $(flo2 \text{ run})_{IV}$
- en IL, en remplaçant dans (iii) c par run' on obtient (iv)

$$\text{iv} \quad [run' \langle IV' \rangle]$$

On remarque ainsi que $Tr(flo2)$ et $Tr(ncouple-pSLF_n)$ n'introduisent aucune modification à $Tr(run_{IV})$. Le $ncouple-pSLF_n$ introduit $flo2$ comme marque-place dans la dérivation de pSLF pour obtenir l'expression finale. Grâce à la manipulation du $ncouple-pSLF_0$ – manipulation qui a des caractéristiques plus simples que celles des manipulations de tous les couples qui ont $F10.n$ comme opération –, ce marque-place va être occupé par *necessarily*, ce qui va permettre d'obtenir

$$\text{v} \quad (... \text{ necessarily runs})_t$$

où ... va être occupé par le sujet. Comme $Tr(ncouple-pSLF_n)$ n'introduit aucune modification dans la traduction et que $Tr(ncouple-pSLF_0) = Tr(< V_9, F_6 >)$, la traduction de (v) restera inchangée par rapport à l'actuelle.

5.3 La généralisation des solutions

La Section 5.2 précédente montre la possibilité de traiter des exemples linguistiques qui sont apparemment hors de portée du système proposé dans PTQ. Les solutions qui y sont données ne sont pas ponctuelles : elles s'inscrivent dans un cadre général qui sera présenté dans les Sections 5.3.2 à 5.3.5, après avoir fait un rappel (Section 5.3.1) des principales permissivités utilisées et de leur rôle, en ce qui concerne aussi bien pSLF que la traduction en form-LI des arbres dérivationnels spécifiés par pSLF.

5.3.1 Panoplie des principales permissivités utilisées et de leur rôle

Les principales permissivités utilisées pour régler les problèmes descriptifs sont récapitulées et commentées par la suite.

Permissivités dans la spécification de pSLF :

- multiplicités des slashes pour distinguer en syntaxe des catégories avec la même traduction ;
- opérations des règles a-catégorielles, et notamment d'effacement et de permutation, et des règles anti-catégorielles ;
- manipulations pour substituer des symboles dans les suites déjà concaténées ;
- expressions élémentaires en V-pSLF sans « correspondance » avec des symboles terminaux dans les suites-LN.

Permissivités dans la traduction :

- variables et types définis en LI et, en particulier, la possibilité d'avoir des variables pour chaque type récursivement définissable ;
- traduction ad hoc en formules lambda d'expressions élémentaires arbitrairement choisies ;
- opérations LI ad hoc dans la traduction des couples-pSLF, ces opérations n'étant pas des CBF-LI; elles relèvent de l' $AdLI$.

L'utilisation d'expressions élémentaires ad hoc, non utilisées dans les suites de la langue naturelle décrite en tant que langage objet, est particulièrement commode pour créer des régularités là où elles n'existent pas dans ce qui est observable du langage objet. Elles permettent de créer des marque-places sur lesquels il sera possible de faire les descriptions.

Les opérations d'effacement et de permutation sont des auxiliaires précieux, d'une part, pour faire disparaître les expressions élémentaires ad hoc dont on vient de parler lorsqu'elles auront rempli leur mission, et d'autre part, pour contourner des exigences de linéarité dans le langage objet. Si dans celui-ci l'entité A va après l'entité B, et si l'on veut, pour conserver l'homomorphisme entre foncteur et argument que ce soit la traduction de A qui précède dans les form-LI la traduction de B, on se donnera une ligne de la dérivation où A est placé avant B dans l'entrée, mais avec une opération qui le permutera avec B dans la sortie.

Enfin, pour conclure avec pSLF, on a déjà remarqué que les manipulations exprimées en langue naturelle ne peuvent pas être assimilées aux expressions d'un formalisme calculable : il s'agit de dire ce que l'on souhaite, mais ce dire est sans contrepartie formelle. On peut ainsi, notamment, faire remplacer n'importe quoi par n'importe quoi d'autre n'importe où.

Du côté permissivités de la traduction, les clefs pour aboutir à ce que l'on veut, tout en respectant les restrictions de la traduction, sont l'utilisation des formules lambda conjointement ou non avec les opérations ad hoc de traduction associées à chaque couple-pSLF. Ces éléments permettent de se libérer de l'utilisation de l'application fonctionnelle sur les types de LI. L'abstraction lambda « à vide » permet d'évacuer des variables introduites dans les traductions. Avec cette gamme de solutions, il devient possible de donner le type désiré au résultat de la traduction de l'opération effectuée.

5.3.2 L'obtention des solutions

Les Exemples 1 à 3 de la Section 4.3, dont les solutions ont été données dans respectivement les Sections 5.2.1, 5.2.2 et 5.2.3, illustrent trois situations qui peuvent se caractériser ainsi :

- l'Exemple 1 (Section 5.2.1) illustre l'exploitation de la multiplicité de slashes ;
- l'Exemple 2 (Section 5.2.2) illustre la problématique de l'ambiguïté catégorielle des expressions terminales ;
- l'Exemple 3 (Section 5.2.3) illustre la problématique des relations d'ordre.

On abordera ces trois points successivement et respectivement dans les trois sections qui suivent.

5.3.2.1 L'exploitation de la multiplicité des slashes

L'Exemple 1 illustre l'utilisation d'objets syntaxiques auxquels on veut faire subir un même sort en traduction en LI. En PTQ cette situation est illustrée par les catégories associées aux noms communs et aux verbes intransitifs. C'est la possibilité d'utiliser une multiplicité

des slashes dans les catégories syntaxiques qui va permettre de traiter les objets ainsi multipliés comme des objets différents en syntaxe, tout en leur donnant une même traduction. On peut ainsi créer une classe d'équivalence. Si l'on a

$$CAT^I/CAT^J, CAT^I//CAT^J, ..., CAT^I/.../{}_nCAT^J$$

on a créé une classe avec des catégories syntaxiquement différentes mais équivalentes par rapport à une même traduction.

C'est ce recours que l'on a utilisé pour faire en sorte que les syntagmes prépositionnels deviennent des modificateurs des noms communs, comme ils sont déjà des expressions verbales qui ont comme catégorie celle des verbes intransitifs, que ces expressions soient élémentaires ou non, et ceci sans modifier la traduction du couple-pSLF $\langle V_{10}, V_7 \rangle$. Dans la Section 5.2.1 on a profité aussi du fait que le ncouple-pSLF_i est, du fait de la permutation, a-catégoriel.

5.3.2.2 Ambiguïté catégorielle des expressions terminales

L'Exemple 2, traité dans la Section 5.2.2, illustre la problématique de l'ambiguïté catégorielle des expressions élémentaires.

Soit la situation générale suivante :

- deux catégories différentes CAT^1 et CAT^2 ;
- un ensemble d'expressions élémentaires telles que :
 - certaines de ces expressions doivent être associées à CAT^1 ,
 - d'autres de ces expressions doivent être associées à CAT^2 ,
 - d'autres de ces expressions doivent être associées à CAT^1 et à CAT^2 .

Dans ces situations on procède ainsi :

- incorporation de $CAT^{2'}$ identique à CAT^2 mais avec un nombre différent de slashes ;
- assignation des expressions non ambiguës à CAT^1 ou à CAT^2 ;
- assignation des expressions à $CAT^{2'}$.

Nous avons déjà remarqué que, dans l'étude des langues naturelles, et notamment dans la tradition philosophique, on a caressé l'idée de traiter l'ambiguïté catégorielle par un recours à l'ellipse. L'idée de fond de cette manière de voir les choses est qu'une expression cesse d'être ambiguë si on la place dans un contexte elliptique, entendant par là qu'on suppose dans ce contexte une entité sans signifiant observable, cette entité elliptique forçant une interprétation unique de l'expression qui serait autrement ambiguë. Ainsi une expression verbale transitive, à laquelle on adjoint un complément d'objet direct elliptique, devient une expression intransitive.

Ce sont les permissivités des Grammaires de Montague qui permettent de concrétiser cette idée. La mécanique du système permet d'obtenir les traductions en accédant aux lignes de la dérivation, à partir desquelles, grâce au pouvoir des opérations, on peut supprimer des symboles. On se donnera ainsi un symbole ad hoc – un symbole *vac* – chaque fois qu'il s'agit de faire en sorte qu'une expression potentiellement ambiguë ne le soit plus dans un contexte elliptique donné, exprimé par ce symbole ad hoc, lequel sera opportunément supprimé par une opération d'effacement.

La traduction en LI du *vac* qui a été placé dans le contexte elliptique sera donnée par une expression lambda ad hoc. Et cette traduction provoquera dans la form-LI les mêmes effets que les effets qui auraient été produits par une expression standard; et on sait que l'on peut disposer d'une telle variable grâce à l'outillage en variables de LI.

C'est exactement ce schéma qui est utilisé pour régler le problème d'ambiguïté de l'expression *eat* (verbe transitif et verbe intransitif). Dans la solution proposée, on a associé l'expression *eat* à la catégorie IV//T, c'est-à-dire, dans le schéma général de ci-dessus, à CAT²'. On a introduit dans V-pSLF un *vac* ad hoc, c'est-à-dire *vac1*. Il va agir comme foncteur d'une expression de catégorie IV//T pour faire en sorte qu'elle devienne de catégorie IV, et il disparaît de la dérivation grâce au pouvoir expressif des opérations qui effacent. En sémantique, *vac1* reçoit une traduction ad hoc, sous forme d'expression lambda, où une variable prend la place de ce qu'aurait été la traduction d'un « vrai » complément d'objet. En fait, c'est une solution dans les lignes de celle proposée dans [Dowty & al. p. 96-97]. On traite de cette manière le *eat* avec complément d'objet direct elliptique, qui devient ainsi intransitif.

Le *eat* avec un « vrai » complément d'objet direct est traité avec l'appareillage habituel qu'on applique aux verbes transitifs; pour ce faire on se donne un nouveau couple avec une règle-pSLF – V5b – calquée sur la règle V5 de PTQ; dans V5b IV//T se substitue à V5 de PTQ. On conserve comme traduction du nouveau couple la traduction du couple homologue (c'est-à-dire <V5, F5>).

On a ainsi :

IV	: expressions verbales toujours intransitives
IV/T	: expressions verbales toujours transitives
IV//T <i>plus vac1</i>	: <i>eat</i> version intransitive en syntaxe mais avec une variable non instanciée en sémantique
IV//T <i>sans vac1</i>	: <i>eat</i> version transitive en syntaxe mais avec la représentation du complément d'objet direct en sémantique

Il est facile de percevoir que cette mécanique permet de traiter une variété de problèmes de sous-catégorisation verbale, et, en particulier :

- verbes transitifs et ditransitifs; p. ex. :

Jean présente Marie

Jean présente Marie à Pierre

- verbes avec compléments optionnels autres que le compléments d'objet direct et indirect

Jean parle

Jean parle de ceci

Jean parle à Marie

Jean parle de ceci à Marie

Le tout consiste à se donner des entrées avec autant de slashes qu'il faudra et à se doter pour chaque cas d'un nouveau *vac* ad hoc.

La question de l'ambiguïté d'assignation catégorielle ne se limite pas à la problématique de la sous-catégorisation verbale. On peut traiter de la même manière ce qu'on a appelé la translation des catégories, p. ex. les expressions qui sont des prépositions et qui sont utilisées comme des adverbes, ou les adjectifs qui sont utilisés comme des noms, situations respectivement illustrées dans (i) et (ii) par la suite :

- i a Jacques viendra après
- b Jacques viendra après la classe

- ii a Jacques a acheté la rectangulaire
- b Jacques a acheté la pièce rectangulaire

Pour traiter (i-a) il faut se donner un vac^T , qui va jouer le rôle d'une expression de catégorie T pour la préposition *après*. On aura ainsi les deux lignes suivantes en pSLF :

IAV/T vac^T
IAV

La traduction de vac^T sera une variable de type $\langle T' \rangle$, et on aura en sémantique une form-LI que l'on pourra gloser « Jacques viendra après quelque chose ».

Pour traiter (ii-a) on procèdera de manière analogue, mais en se donnant un vac^{CN} que l'on introduira dans une ligne de la dérivation afin que l'adjectif – que l'on suppose foncteur de catégorie CN/CN puisse trouver son argument à droite, et on fera disparaître ce vac^{CN} pour ne pas le voir dans la suite terminale. Parallèlement, en sémantique, on dispose d'une variable qui aura le type $\langle CN' \rangle$. La traduction finale de la form-LI résultante pourra être glosée par « Jacques a acheté quelque chose qui est rectangulaire ».

5.3.2.3 Relations d'ordre

L'Exemple 3, traité dans la Section 5.2.3, illustre une partie de la problématique des relations d'ordre entre symboles terminaux dans les suites-LN. On sait que les relations d'ordre dans les langues naturelles diffèrent de langue à langue, mais en général ne sont ni complètement libres ni contraintes à tel point qu'on puisse en rendre compte moyennant la

concaténation par adjacence : si, p. ex., en français, en anglais et en espagnol on peut rendre compte par adjacence de la plupart des entités dans le syntagme nominal noyau, dans ces mêmes langues l'emplacement des adverbes est beaucoup plus libre. Les adverbes en effet ont la fastidieuse particularité de se déplacer, en apparence capricieusement, dans la phrase. Et ceci est particulièrement gênant pour les systèmes contraints, comme p. ex. les grammaires de niveau 2, ou pour l'application fonctionnelle, qui doivent évaluer des règles en concaténant par adjacence et sans permutations possibles. Dans ces problèmes des relations d'ordre on n'envisage pas la problématique des éléments arbitrairement éloignés qui pose encore d'autres problèmes.

La possibilité d'utiliser des opérations permutantes et, surtout, d'inclure dans les descriptions des couples-pSLF des manipulations, règlent en PTQ, côté syntaxe, le problème de l'emplacement d'entités à linéarité « laxiste », c'est-à-dire avec des relations d'ordre qui ne sont pas complètement libres, mais qui ont un très grand potentiel des déplacements possibles. La possibilité d'utiliser les expressions lambda en LI règle la question d'instancier une variable correctement placée dans la form-LI par une constante qui sera concaténée ultérieurement à l'emplacement de la variable.

Le schéma général pour traiter cette problématique consiste donc en ceci :

- Se donner des symboles *flo* de deux classes différentes :
 - les *flo1* introduisent la variable adéquate dans la form-LI et vont disparaître de la syntaxe par application d'une opération d'effacement ;
 - les *flo2* introduisent un marque-place dans la suite concaténée en pSLF au moment de leur concaténation, et leur traduction en sémantique fait que rien ne change à la sémantique (form-LI) qui était associée à ce qui avait déjà été concaténé.
- La variable introduite par les *flo1* sera instanciée par la constante dans laquelle aura été traduite l'expression qui sera attestée en définitive dans la suite terminale lorsque cette expression sera concaténée; ceci grâce au pouvoir expressif des expressions lambda.
- Les marque-places laissés en arrière par les *flo2* seront occupés, grâce à une manipulation, par l'expression qui sera finalement attestée dans la suite terminale finale, laquelle, au moment de sa concaténation, disparaîtra ainsi, mais non sans avoir incorporé sa sémantique à la form-LI.

On peut ainsi traiter A et B, les deux situations générales suivantes. Supposons une suite avec deux positions, la position *j* étant à droite de la position *i* :

... *i* ... *j* ...

Situation A : une expression doit être placée en *i* et en *j*, mais, dans les deux cas, elle apporte à la form-LI la sémantique résultante de la position *j*; cette situation sera traitée au moyen des *flo1*.

Situation B : une expression doit être placée en *i* et en *j*, mais dans les deux cas, elle apporte à la form-LI la sémantique résultante de la position *i*; cette situation sera traitée au moyen des *flo2*.

6 Éléments pour une réévaluation de la Grammaire de Montague et de ses interprétations

La Grammaire de Montague a été l'un des grands événements dans l'histoire de la linguistique, science du langage. Elle a été examinée ici d'un point de vue censé être éclairant, mais qui est sans doute partiel, à savoir sur son empiricité éventuelle pour étudier les langues naturelles en tant que langage objet. L'utilisation de la LI que Montague propose pour traiter maints puzzles sémantiques exhibés par les langues naturelles n'a en rien été abordée, alors que c'est l'un de ses apports fondamentaux (cf. une discussion approfondie de ce point dans le dernier chapitre de [Chambreuil & Pariente 90]). On s'est cantonné à analyser ses descriptions syntaxiques et leurs rapports avec leur traduction ou leur interprétation.

Le parcours de l'analyse a nécessité une clarification du domaine, quelque peu perturbé par l'utilisation de concepts, notions ou interprétations de type auberge espagnole d'autrefois, qui, se référant en vrac à Montague, ont introduit des glissades interprétatives dangereuses, en particulier celles induites par les notions de compositionnalité^{aea} ou de compositionnalité^{Pi}, avec, en accompagnement, la notion de parallélisme syntaxe-sémantique.

Comme exemple de cette absence de clarté, soit une des interprétations courantes de la compositionnalité, que nous classons en compositionnalité^{aea}, [Partee 97 p. 21] signale que :

This way of implementing the compositionality requirement in terms of corresponding recursive rule specifications of syntax and semantics has been dubbed by Bach (1976) the requirement of "rule-by-rule interpretation" and it is the form in which Montague grammars are most commonly instantiated.

Le trouble vient lorsque la compositionnalité, toujours selon des auteurs prestigieux, est située ailleurs. On sait que GPSG satisfait « the requirement of « rule-by-rule interpretation », alors que dans le même temps [Gazdar & al. p. 187] met en doute la possibilité de définir l'ensemble de types de LI requis selon PTQ « as a homomorphism from [...] syntactic categories ». Ainsi selon que la compositionnalité et/ou l'homomorphisme sont situés sur la notion de « règle à règle » ou sur celle des contraintes sur la traduction des catégories en types, GPSG satisfait ou non la compositionnalité¹². C'est ce type même

¹² Dans [Janssen 91, p.431] on donne un satisfecit à la compositionnalité dans GPSG, et ceci donc malgré l'absence systématique de traduction des catégories syntaxiques en types de LI.

de confusions qui nous ont conduit à durcir l'analyse en fonction de ce que nous croyons être effectivement dit ou n'être pas dit aussi bien dans UG que dans PTQ.

Les présentations intuitives de notions, concepts, idées, etc. sont certes indispensables pour comprendre de quoi on parle. Elles deviennent dangereuses lorsqu'on construit à partir d'elles des dérivées qui n'ont pas lieu d'être à partir de formulations plus rigoureuses.

Revenons aux présentations intuitives de la compositionnalité de la Section 3.2, c'est-à-dire à la compositionnalité^{Pi}. Nous croyons que le dernier alinéa de la citation de [Chambreuil & Pariente 90 p. 64] de la Section 3.2, rappelée ici dans (i), devrait être reformulée comme dans (ii).

(i) Ainsi l'expression langagière « jean marche lentement dans le parc » ne sera plus associée à une expression élémentaire [de la logique des propositions] comme c'était le cas en logique propositionnelle.

Elle ne sera pas non plus associée à une expression [de la logique des prédicats]. L'expression langagière sera obtenue elle-même comme expression du langage formel, construite à partir des expressions élémentaires : « Jean », « marcher », « lentement », « dans », « parc », et aussi « un » (p. 64).

(ii) L'expression langagière, *Jean marche lentement dans le parc*, sera obtenue elle-même comme suite terminale d'une dérivation spécifiée par le langage formel [pSLF], cette dérivation ayant comme suite basique – suite initiale de la dérivation – la suite

« Jean » « dans » « le » « park » « lentement » « marcher »

et en utilisant des règles dont :

- quelques-unes ont le pouvoir expressif de l'application fonctionnelle des grammaires catégorielles
- d'autres ont le pouvoir expressif des règles de niveau 0 dans la hiérarchie chomskyenne
- d'autres enfin ne sont pas explicitées du tout, seulement suggérées dans une expression en langue naturelle standard.

L'énorme et cruciale différence entre la compositionnalité^{Pi} et ce qui est dit dans UG et pratiqué dans PTQ, vient de ce qui est souligné par nous en gras dans la citation (déjà présentée dans la Section 3.2) suivante de [Dowty et al. 81 p. 179] :

...by means of recursive definitions specifying how complex phrases are to be formed out of simpler ones. This procedure makes for a "Fregean" semantic treatment of English [...] the syntactic rules [de pSLF] will not only concatenate input expressions, **possibly with other symbols added as well, but may also perform transformation-like manipulations of the inputs.**

Ce qui est exprimé en gras dans la citation précédente – qui est en fait ce qui permet l'adjonction et la suppression des symboles de la dérivation dans PTQ – est strictement

absent et insoupçonnable à la lecture de la présentation intuitive du même point dans les pages liminaires :

Frege's Principle [...] places fairly severe constraints on the systems of syntactic semantic rules we can construct. Note, for example, that "the parts" referred to in the statement of Frege's Principle must be syntactic constituents of the expression in question [Dowty et al. 81 p. 8-9].

Il est bien probable que les présentations intuitives de la compositionnalité montaguienne aient été à l'origine de certaines idées sur un type de lien entre Grammaire de Montague et traitement automatique, lien que, tel qu'il est présenté dans la citation suivante, nous ne percevons pas.

La méthode de calcul induite par le principe de compositionnalité est par ailleurs incrémentale et monotone. On peut élaborer une structure complexe à partir des structures les plus élémentaires, en construisant pas à pas des structures de complexité croissante (incrémentalité). Les structures sémantiques sont construites une fois pour toutes : une sous-structure ne peut être remise en cause lors de son intégration dans une structure complexe (monotonie). Ces deux propriétés sont importantes parce qu'elles caractérisent un type d'opérations que l'on sait modéliser, pour lesquelles on dispose de modèles de calcul, l'unification notamment [Nazarenko 98 p. 4].

Il se peut que la compositionnalité^{Pi} puisse induire une méthode de calcul : en fin de compte le propre des notions intuitives est que l'on peut les interpréter flexiblement. Mais nous ne percevons pas cet effet supposé de la compositionnalité en fonction de ce qui est dit dans UG, car aucune contrainte n'y est introduite sur l'algèbre qui va spécifier la syntaxe de LN. Nous ne percevons pas non plus cet effet supposé de la compositionnalité dans la pratique de PTQ, où des manipulations vont même substituer des symboles déjà introduits dans la dérivation (cf. p. ex. le traitement de l'anaphore et de la quantification)¹³.

Dans cette révision des interprétations intuitives de la Grammaire de Montague, il est difficile de ne pas s'attarder sur des affirmations portant sur la relation de celle-ci avec les

¹³ [Janssen 97 p. 456] évoque un théorème [Zadrozny 94] prouvant que n'importe quelle sémantique peut être traitée de manière compositionnelle mais qui ne s'applique qu'à une syntaxe relevant de la compositionnalité^{Pi} (une syntaxe où « there is concatenation of visible parts »), celle-ci étant considérée « very attractive » par [Janssen 97]; mais, justement, la syntaxe de PTQ ne relève pas de ce type. Ceci a sans doute posé problème à l'analyse automatique de PTQ ; cf. sur cette question les tentatives de [Friedman & Warren 78], [Warren & Friedman 82], [Landsbergen 83]. Il est difficile de se rendre compte de quelles sont les parties de PTQ que ces tentatives peuvent effectivement analyser automatiquement. Dans la présentation de ce type de travaux (en particulier les deux premiers ci-avant mentionnés), [Chambreuil 89 p. 175] évoque la « réalisation informatique d'aspects du système de Montague, ou sur l'utilisation de ce système (ou de notions inspirées de ce système) pour la réalisation d'une tâche liée à l'activité langagière ». En somme, ce n'est pas PTQ qui est traité, mais un aspect partiel mal délimité. [Landsbergen 83] considère que le système de UG « has to be adjusted in such a way that for the grammars within the modified framework a parser can be defined ». À cet effet il définit des « M-grammars » dont le domaine « in which the M-rules operate, i.e. the set of labelled bracketings that are allowed as possible inputs and outputs of these rules, is defined separately, by means of a context-free grammar. The language defined by the M-grammar is a subset of the language defined by this context-free grammar ». Il est évident qu'avec ces caractéristiques, les M-grammars ne peuvent traiter qu'une partie très restreinte de PTQ. Tout se passe donc d'une manière tout à fait analogue à ce qu'on observe par rapport à la compositionnalité^U en relation à la syntaxe de PTQ : si l'on se donne des syntaxes avec des restrictions, il faut laisser tomber la compositionnalité^U pour se donner d'autres compositionnalités, moins exigeantes ; sur ce point cf. la Note complémentaire 1.

grammaires catégorielles. Soit les trois citations suivantes, qui relèvent toutes de [Partee 97] :

- i [Un certain type de grammaires] and the various categorial grammar frameworks [...] are among the clearest examples of "linguists' grammars" that are more or less consistent with the requirements of Montague's UG (p. 19).
- ii Montague[...] used a modified categorial grammar for PTQ (p. 30).
- iii [...]it is probably no accident that PTQ's "Rules of Functional Application" cover virtually all of the basic grammatical relations instantiated in PTQ (p. 48).

Ce qui est dit dans (i) ne nous paraît pas significativement adéquat, car l'absence des contraintes sur l'algèbre de UG fait que n'importe quel « modèle » linguistique, pourvu qu'il soit formulable comme une algèbre, satisfasse ce qui est demandé par UG; en fait, de la citation de [Dowty & al. 81 p. 179] se déduit que la grammaire transformationnelle, dans l'une de ses expressions, était aussi compatible¹⁴; cf. par ailleurs l'adaptation de [Bach 79]. Et il nous est impossible de partager l'affirmation de (ii), et, tout particulièrement, celle de (iii). Pour étayer notre disconformité, le lecteur est invité à se rapporter au tableau des couples-pSLF présenté dans la Section 4.1.3 : sur 24 couples-pSLF il n'y en a que 6 qui peuvent s'exprimer par une grammaire catégorielle sans manipulations et en utilisant l'application fonctionnelle. Et les descriptions les plus intéressantes (p. ex. coordination, relatives, anaphore) échappent toutes à l'application fonctionnelle. Certes, on peut définir comme on veut l'idée de « modification », mais c'est tirer un peu trop sur la définition que de dire que Montague utilise dans PTQ une version modifiée des grammaires catégorielles, même s'il est vrai que quelques couples-pSLF peuvent s'exprimer par l'application fonctionnelle des grammaires catégorielles.

Pour compléter cette analyse critique des présentations plus ou moins intuitives de Montague on revient sur le *p* de *pSLF*. Dans ce document nous utilisons *manipulation(s)* pour dénoter tout ce qui dans les CBF-pSLF est exprimé en langue naturelle. Ceci n'est pas l'usage courant, où très généralement on utilise *règle* pour parler aussi bien de ce qui est formel et de ce qui relève d'un souhait de formalisation à venir. Ces usages, pour nous malheureux, sont illustrés par les quatre citations suivantes¹⁵ :

- i [...] nous supposons que l'opération F_4 est remplacée par une opération $*F_4$ et la règle V_4 [...] par une règle $*V_4$ [...] que pour tout n de N , $F_{10,n}$ est remplacée par $*F_{10,N}$, et que les règles $V_{16,n}$, $V_{17,n}$, $V_{18,n}$ sont remplacées respectivement par : $*V_{16,n}$, $*V_{17,n}$, $*V_{18,n}$ [Chambreuil 89 p. 67-68].
- ii As an example of a syntactic rule, consider S_4 [cf. le couple pSLF $\langle V_4, F_4 \rangle$], the so called "subject-predicate" rule [...] [Dowty & al. 81 p. 191].
- iii It was noted above that PTQ has three Quantifying In rule [déjà citées dans (i)] " [Partee 97 p. 37].
- iv The syntactic rule for restrictive relative clause formation in PTQ [cf. le couple pSLF $\langle V_{3,n}, F_{3,n} \rangle$] [...] [Janssen 97 p. 436].

¹⁴ Mais vraisemblablement pas d'autres formulations plus récentes, cf. [Janssen 97 p. 431].

¹⁵ Ces usages continuent celui instauré dans PTQ.

Les citations de ce type pourraient être multipliées : dans tout les cas ce qu'on nomme *règles* (anglais *rules*), incorporent des pseudo-opérations, que nous avons nommées *manipulations*. Ce sont ces manipulations qui effectuent la partie la plus difficile des opérations à formuler un jour. On sait en TALN que traiter des relatives, de la coordination, de l'anaphore relève de l'exploit, largement encore à accomplir. Or, pour traiter de ces constructions, les manipulations sont systématiquement requises. Et on peut dire que ces manipulations sont basiquement mal conçues : elles ont été imaginées dans l'espoir qu'une exploration linéaire plate de la suite terminale, sans tenir compte des parenthésisations internes à cette suite, pouvait aboutir à la détection du bon élément à remplacer ou à modifier¹⁶. Or si l'on a appris quelque chose dans le traitement automatique des langues naturelles, c'est que cette possibilité n'existe pas. C'est sur ces points concrets et identifiables que l'on peut mesurer la distance entre les idées directrices de Montague et leur application à l'étude des objets que l'on veut étudier, à savoir les langues naturelles. C'est sur ce type d'observations qu'on comprend la signification de l'affirmation suivante de [Partee 97 p. 48] et les limitations qu'elle induit, affirmation, à laquelle, cette fois, nous adhérons sans réserve :

[...] Montague put no constraints on the syntactic operations that could be used in a syntactic rule, not even a constraint on computability.

Dans les Sections précédentes, pour évaluer l'empiricité de UG et de PTQ, nous avons porté l'analyse sur le formalisme utilisé dans ces travaux – ou sur l'absence de formalisme – plutôt que sur les données que ces formalismes sont censés expliquer ou décrire. En faisant cela nous nous sommes pliés à ce qui est très frappant dans les écrits de Montague, à savoir une absence de perspective observationnelle systématique.

Nous avons remarqué (cf. Section 5.3) l'utilisation de symboles vides et des opérations d'effacement pour traiter, parmi d'autres, l'assignation catégorielle ambiguë des suites élémentaires. Nous interprétons ces procédés comme des outils permettant la mise au pas de ce qui, moyennant des procédés généraux, est directement observable en langue naturelle. Montague rejoint ainsi, sur ce point, la tradition philosophique.

Nous avons par ailleurs remarqué l'utilisation du vocabulaire *Ve3-pSLF* (avec les éléments de type *he0*). Or, ces formes, si on les laisse, défigurent dans les limites d'aucune restriction générale ce qui est directement observable (suites des mots graphiques). Cela a soulevé des « interrogations » dans [Chambreuil 89 p. 53]. À notre avis, ces « interrogations » ne peuvent se résoudre que par l'une des deux réponses suivantes, les deux introduisant des bémols sérieux :

- Si la description porte sur des suites avec des éléments de *Ve3-pSLF*, on n'étudie pas l'anglais, mais un anglais mis au pas, et la mise au pas se fait dans un cadre irréstrict.

¹⁶ [Morrill 94 chap 1] signale la faiblesse formelle de la syntaxe de PTQ; en ce qui concerne l'impossibilité d'une exploration plate dans les manipulations, cf. [Morrill 94 p.13 note 4].

- Si l'on se donne dans R une opération de substitution de ces éléments par des pronoms de l'anglais, on obtient des résultats inadéquats (mauvaise résolution de l'anaphore, cf. la discussion de ce point dans la Section 4.3).

Revenons à l'Idée 1 de Montague, citée dans la Section 2 en discours rapporté à partir de [Chambreuil & Pariente 90 p. 19] et représentée ici dans sa version originale :

There is in my opinion no important theoretical difference between natural languages and the artificial languages of logicians; indeed, I consider it possible to comprehend the syntax and semantics of both kinds of languages within a single natural and mathematically precise theory (UG, p. 222).

Nous avons exposé l'argumentation qui nous a conduit à refuser d'interpréter cette affirmation comme hypothèse empirique et, par ailleurs, nous nous sommes situé par rapport aux interprétations courantes et/ou intuitives de la Grammaire de Montague. Nous introduisons ici une nouvelle limitation, cette fois sur le plan formel :

L'intégration de la syntaxe et la sémantique des deux types de langages – les langues naturelles et les langages artificiels des logiciens – dans une unique théorie mathématique, naturelle et précise, n'a jamais vu le jour jusqu'ici : cette théorie mathématique n'a jamais été formulée.

Cette limitation se fonde principalement sur les constats suivants :

- Pour pouvoir aboutir à l'interprétation induite, l'ensemble de définitions (cf. leur rappel dans la Section 3.2.1) nécessite des conjectures présentés dans UG sous l'intitulé de *Remarks*.
- Pour prouver ces conjectures, il faut passer par la preuve de l'existence de la classe K d'opérations polynomiales (cf. la définition (ii) de Section 3.2.1).
- Pour prouver l'existence de la classe K d'opérations polynomiales, il faut passer par une axiomatisation particulière des théories des ensembles.

Sur le dernier point, citons encore une fois UG p. 224 note 4 :

For those interested in set-theoretic technicalities I might point out that in this characterization [caractérisation des opérations polynomiales de K] the word 'class' is used deliberately rather than 'set'. In any of the usual axiomatic formulations of set theory one can prove that there is no *set* K satisfying (1)-(4), but in those formulations that recognize proper classes in addition to sets one can prove that there *is* a proper class satisfying those conditions.

Or, autant qu'on sache, personne n'a pris la peine de prouver les quatre Remarques de UG, ni d'explicitier l'axiomatisation de la théorie des ensembles permettant de prouver l'existence de la classe K des opérations polynomiales ni, à fortiori, de prouver l'existence de cette classe. En résumé donc, syntaxe et sémantique des langues naturelles et des logiques relèvent d'une même théorie mathématique... qui n'a jamais été formulée.

Que reste-t-il de l'Idée 1 et de sa concrétisation dans la compositionnalité?

À notre avis, nous avons une espèce de commandement méthodologique : *l'homomorphisme syntaxe-sémantique tu construiras*. Pour ceux qui y croient (p. ex. [Nazarenko 98]), c'est bien de le satisfaire.

Ainsi, p. ex. [Janssen 97] affirme à plusieurs reprises le rôle méthodologique de la compositionnalité :

[The principle of compositionality] should not be considered an empirically verifiable restriction, but a methodological principle that describes how a system for syntax and semantics should be designed (p. 419).

The challenge of compositional semantics is not to prove the existence of such a semantics, but to obtain one [...] Compositionality is not an empirical principle, but a methodological one (p. 457).

Pour [Janssen 97] il s'agit donc de faire du compositionnel, car de toute manière, pour ce faire « enough methods are available » (p. 441), et il en passe en revue plusieurs.

Nous ne nous attardons pas sur ces méthodes, mais nous remarquons deux points. L'un est le recours à l'assignation catégoriellement ambiguë des expressions élémentaires. Ainsi p. ex. la solution donnée à *any* :

We conclude that the any-examples can be dealt with in a compositional way by distinguishing two *any*'s, with one or both readings eliminated when incompatible with the surrounding context [Janssen 97 p. 440].

Or pour nous cette méthode est inapplicable pour aboutir à l'homomorphisme-u exigé par Montague dans UG, cf. la discussion de ce point dans Ob. 3 de la Section 3.2.1 et dans l'Annexe.

L'autre est l'utilisation des postulats de signification. Ces postulats effacent, parmi d'autres fonctions, les lectures surgénérées; cf. p. ex. leur utilisation pour effacer les lectures surgénérées pour des constructions du type *manger un sandwich*, qui n'ont pas besoin de toutes celles qu'il faut associer à *chercher une licorne*. Et ici le problème est le suivant : ou bien on n'efface pas et en surgénérant on obtient des résultats observationnellement inadéquats, ou bien on efface et on utilise alors une opération qui, certes, n'est pas dans l'algèbre de la syntaxe, mais qui vient aider celles qui y sont, et dont on voit mal le corrélat dans l'algèbre dérivée de LI (cf. l'utilisation de ces postulats dans la formulation VIII de la Section 7.2 comme un des points distinguant les langues naturelles des langages logiques)¹⁷.

¹⁷ Les postulats de signification introduits par Carnap sont utiles pour traiter des questions de synonymie et de paraphrase. Or ce n'est pas ce rôle qu'ils jouent ici (cf. la discussion de ce point dans [Dowty & al. 81 p. 224-225]), où ils agissent en tant que filtres de la surgénération provoquée par la technique de la pire hypothèse adoptée par Montague pour traduire les catégories syntaxiques en types de la logique, en insérant systématiquement l'entité *s*.

Même en laissant de côté l'efficacité et/ou la possibilité d'utiliser les méthodes d'obtention des homomorphisme-u proposées, leur utilisation doit sûrement contenter davantage ceux qui sont plus intéressés par la satisfaction de contraintes méthodologiques qu'à ceux qui s'intéressent aux descriptions effectives et vérifiables par des procédés de calculs, eux aussi effectifs. Sur ce point crucial, nous adoptons donc le point de vue de Hintikka et Pelletier, discutés de manière non convaincante dans [Janssen 97]. Remarquons simplement en passant que le fait de vouloir aboutir à une compositionnalité formelle, comme celle proposée dans [Janssen 97] exige de se donner des objets en syntaxe et/ou sémantique plus complexes, comme le montrent les discussions avec Pelletier et Hintikka, et que cette compositionnalité forcément trouvée par obéissance à un commandement méthodologique n'est jamais mise en perspective par rapport à l'Idée 1 de Montague, dont la clef pour être vérifiée est l'existence empiriquement testable et formellement possible de l'interprétation induite¹⁸.

D'après nous, il ne semble pas qu'il soit intéressant de suivre ce commandement pour deux raisons, qui viennent de la méthodologie générale des sciences empiriques :

- Aucun principe méthodologique ne garantit par lui-même la qualité du produit obtenu en le respectant.
- Si l'homomorphisme-u se situait comme une manifestation d'une idée générale consistant à essayer de trouver un justificatif partiel pour chaque hypothèse d'un ensemble d'hypothèses, on retrouverait l'illusion de pouvoir partialiser la justification de celles-ci.

Ces deux raisons sont étayées par les citations suivantes de [Bunge 69] correspondant respectivement aux deux raisons précédentes.

[Bunge 69 p. 26] rappelle l'échec des philosophes qui, depuis Bacon et Descartes, ont eu la prétention de connaître les règles infallibles de la direction à imprimer au processus de recherche. Et après avoir mentionné plusieurs principes méthodologiques, il insiste sur le fait que les règles de la méthode scientifique sont très loin d'être infallibles, en ajoutant [c'est nous qui traduisons] :

Nous ne devons pas espérer que les règles de la méthode scientifique viennent se substituer à l'intelligence comme résultat d'un simple dressage subi avec patience. La capacité de formuler des questions subtiles et fécondes, celle de bâtir des théories fortes et profondes, et celle d'arbitrer entre des tests empiriquement fins et originaux, ces capacités ne sont pas des activités orientées par des règles; si elles l'étaient, comme certains philosophes l'ont supposé, tout le monde pourrait mener à bien des recherches scientifiques [...] [Bunge 69 p. 27].

¹⁸ Cette observation vaut également pour d'autres ressources utilisées pour traduire la syntaxe. Ainsi p. ex [Partee 97 p. 19 note 12] indique que « Actually, there is a way of respecting the homomorphism requirement while working with semantically ambiguous expressions. It is standard in mathematics to turn a (one-many) relation into a function by making it a set-valued function ». Nous n'avons évidemment rien contre cet usage permettant de transformer des relations en fonctions. L'interrogation porte sur la possibilité d'utiliser cette manière de s'en tirer si l'on veut traduire la syntaxe dans la logique LI de Montague et obtenir l'interprétation induite sur l'univers de dénotation, tel que demandé dans UG.

Tester (*constraster* dans la terminologie de Bunge) un ensemble d'hypothèses, processus « intriqué » d'après [Bunge 69 p. 877], n'exige pas « que toute hypothèse soit *directement* testable : cette exigence, présentée par l'empirisme, rendrait impossible la théorie scientifique ». Par ailleurs :

Ce ne sont pas tous les composants d'une théorie factuelle qui peuvent individuellement avoir un corrélat réel propre [...] En plus, il est douteux qu'on puisse leur assigner des contreparties matérielles individuelles. Ce qui est plutôt normal c'est que ce soit l'image théorique *total*, le *tout* du système [...] qui ait telle contrepartie dans le réel [...] En le disant brièvement : la correspondance entre les modèles théoriques et ses corrélats n'est pas ponctuelle, mais une correspondance de *système* à *système* [Bunge 69 p. 420-421, et cf. aussi p. 304-307].

Il semble assez plausible que le commandement méthodologique sur la compositionnalité puisse remonter à la croyance que le respect des principes méthodologiques produit de « bons » résultats. Dès que cette illusion tombe, la compulsion du commandement disparaît aussi. Il est moins sûr que l'on puisse faire remonter la filiation du commandement méthodologique sur l'homomorphisme à l'idée générale qui veut la partialisation de la justification des hypothèses, même si les deux attitudes ont une vision commune de la relation entre formules dans un langage et partialité dans un réel. En tout cas c'est un point clef de la méthodologie adoptée par 5P de ne pas essayer de justifier partiellement le système des hypothèses descriptives ou explicatives qu'on formule et ceci en parallèle à la non-obéissance au commandement méthodologique sur la compositionnalité, si celle-ci était conçue en tant qu'homomorphisme entre syntaxe de la langue naturelle et univers de dénotation¹⁹.

7 Empiricité en linguistique et homomorphisme syntaxe-univers de dénotation

Cette Section esquisse la possibilité d'une utilisation empirique des morphismes algébriques. Dans la Section 7.1 on montre ce que peut être dans une science empirique – la cristallographie – cette utilisation. Nous croyons que, telle qu'elle se pratique dans les sciences empiriques, elle n'est pas directement transposable à la linguistique. Cependant, il se pourrait que l'homomorphisme visé par Montague puisse être utilisé pour caractériser empiriquement les langues naturelles. Cette possibilité est explorée dans la Section 7.2.

¹⁹ Ceci en accord complet sur ce qui est dit à propos de la méthodologie et, plus généralement, sur le schéma des sciences de l'empirie dans [Sokal & Bricmont chap. 3]; on n'est donc pas en train de dire que l'on est d'accord avec [Feyerabend 79], position largement tempérée dans les éditions postérieures de cet ouvrage et dans [Feyerabend 96 chap. 12].

7.1 Du sel de cuisine et du chlorure de potasse ou de l'utilisation empirique des morphismes algébriques

Le sel de cuisine (ou *halite* NaCl) et le chlorure de potasse (ou *silvine* KCl) sont des espèces minérales directement observables. Elle ont des propriétés macroscopiques différentes. Le sel est indispensable à la diète humaine, tandis que le chlorure de potasse n'est pas comestible. Les deux sont constitués d'atomes différents.

Cependant sel et chlorure de potasse, sur le plan cristallographique, ne sont pas homomorphiques, mais sont liés par une relation algébrique encore plus forte : ils sont isomorphiques. Leurs cristaux, aussi bien ceux de l'un que ceux de l'autre, ont une structure dite de cube. Plus exactement, ils appartiennent à ce que les géologues nomment un même groupe spatial (en l'occurrence le $Fm3m$), un groupe spatial étant défini par des opérations de symétrie définies en algèbre vectoriel. Cette appartenance, qui les rend isomorphiques, n'implique pas l'identité des paramètres cristallins mais une constance des rapports entre les entités de chaque espèce occupant les positions dans les cubes des cristaux respectifs²⁰.

Cette structure cubique dans les cristaux du sel et du chlorure de potasse n'est pas le résultat de l'imposition à l'un de la structure formelle de l'autre. Elle n'est pas le résultat de l'obéissance à un commandement méthodologique du type *homomorphismes tu construiras*. On a découvert dans le réel, observé d'une certaine manière, que des choses observées comme différentes selon des critères établis, avaient, à un certain niveau, abstrait mais associable de manière explicite à ce qui peut être observé, une structure algébrique commune. Qu'il soit clair :

- que cette structure n'est pas une propriété directement accessible par les sens humains; on n'est donc pas en train de dire qu'une chose n'est pas scientifique parce qu'elle porte sur des objets non directement observables ;
- que cette structure résulte d'un appareil conceptuel de cette science du réel qu'est la cristallographie : on n'obtient pas les cristaux n'importe comment ; ce n'est pas une structure imposée par un système de description, mais une construction hypothétique qu'on teste ;
- que toutes les conditions et de calcul et d'observation sont là pour pouvoir tester, par rapport à l'observation, que le cristaux du sel ont bien une structure cubique, que les cristaux de chlorure de potasse ont bien une structure cubique, que les rapports entre

²⁰ Nous remercions le professeur Ángel Fernández, de l'Université de Limoges, d'avoir bien accepté de nous écouter et de nous comprendre lorsque nous nous sommes adressé à lui en quête des données précises sur l'utilisation des morphismes algébriques dans les sciences empiriques. Tous les détails techniques présentés dans le texte viennent de lui ou de la bibliographie à laquelle il nous a fait accéder et toute erreur n'est imputable qu'à une incompréhension de notre part. La description, surprenante pour le néophyte, des structures cubiques isomorphes de la halite et de la silvine peut être trouvée dans Alain Baronet, *Minéralogie*, (Paris, Dunod, 1988). Sur la datation de l'utilisation systématique des morphismes en cristallographie, cf. Pierre Lapadu-Hargues, *Précis de Minéralogie* (Paris, Masson, 1954). À remarquer qu'en cristallographie on réserve le terme d'*isomorphisme* pour des relations entre structures des cristaux qui, en plus d'être isomorphiques dans le sens algébrique, présentent des éléments particuliers dans ces structures. Pour les deux espèces utilisées dans l'exemple – le sel de cuisine (ou *halite* NaCl) et le chlorure de potasse (ou *silvine* KCl) –, on parle en cristallographie d'*isotypisme*, ce qui ne change évidemment rien au fond de la discussion.

éléments spatialement situés dans ces deux structures sont les mêmes dans les deux, alors que chaque structure cubique reste indépendante l'une de l'autre, dans ce sens que chacune a ses propres conditions d'adéquation; ce n'est pas la religion de chaque cristallographe qui va être la compulsion déterminante, via la satisfaction d'un commandement, pour justifier la trouvaille du morphisme. Si l'on décrivait la structure cristallographique du sel par une structure autre que cubique, cette structure serait observationnellement inadéquate et ceci indépendamment de l'adéquation observationnelle de la structure cubique du chlorure de potasse. *Mutatis mutandis* cela est vrai aussi pour le chlorure de potasse.

Ce n'est pas vrai que l'on a toujours su que sel et chlorure de potasse avaient la même structure cristallographique. En fait, l'observation systématique des cristaux en fonction des morphismes algébriques est à situer dans le premier tiers du XIX^e siècle. La connaissance a augmenté lorsqu'on a pu établir les hypothèses sur les morphismes entre les structures des cristaux d'espèces minérales différentes. Et si l'on a pu avancer dans la connaissance, c'est parce qu'on a fait une affirmation qui aurait pu être fausse, et qui reste, en quelque sorte, sur le fil du rasoir : rien ne garantit qu'un jour elle ne deviendra pas fausse²¹. Mais le point essentiel en est que, selon l'appareillage conceptuel et observationnel de la cristallographie, n'importe quel minéral ne peut pas être déclaré homomorphique ou isomorphique avec n'importe quel autre minéral.

Voilà exactement ce qu'on n'a pas fait en philosophie du langage avec la mise au pas de la langue naturelle. Voilà ce que Montague n'a pas fait avec ses Idées 1 à 3, et son homomorphisme entre structures algébriques de la langue naturelle et structures algébriques de LI et/ou de l'univers de dénotation. Voilà ce que Janssen n'a fait pas en appelant à se conformer au commandement *homomorphisme syntaxe-sémantique tu construiras*.

7.2 L'homomorphisme syntaxe-univers de dénotation et sa possible utilisation empirique ou la naturalité de la langue naturelle

En partant de l'examen de l'Idée 1 de Montague sur l'inexistence des différences théoriques fondamentales entre langue naturelle et langages logiques, nous avons développé une argumentation aboutissant à la conclusion que, si cette absence de différence se situait dans le fait que les deux types de langages pouvaient être associés par un homomorphisme- u à l'univers – dans les cas des langues naturelles via un langage intermédiaire dont en principe on peut s'en passer – elle ne pouvait pas être interprétée empiriquement comme une propriété des langues naturelles, étant donné :

- l'absence de restrictions sur l'algèbre de la syntaxe des langues naturelles ;
- l'absence de caractérisation quelque peu précise des types de données observationnelles sémantiques dont la description linguistique doit rendre compte ;

²¹ En fait les structures cristallines ont pu être vérifiées après l'application de la diffraction des rayons X à l'étude des cristaux.

- l'absence de caractérisation de la relation R qui relie l'objet formellement décrit avec l'objet directement observé.

Grâce à toutes ces absences il est impossible de ne pas trouver une algèbre pour la syntaxe de la langue naturelle telle qu'elle puisse être associée par l'homomorphisme-u à l'univers de dénotation. Vu selon cette perspective, l'homomorphisme-u, pièce clef pour concrétiser l'Idée 1 de Montague, ne peut rester qu'en tant que commandement méthodologique, dont la satisfaction ne peut satisfaire que ceux qui y croient.

Et pourtant, sans changer d'opinion sur les conclusions précédentes, nous allons essayer, dans le restant de cette Section 7.2, de faire un usage empirique de l'homomorphisme-u, c'est-à-dire de l'utiliser pour exprimer des caractéristiques empiriques des langues naturelles.

Pour établir ce renversement dans l'utilisation de l'homomorphisme-u, nous devons au préalable revisiter le cadre quelque peu polémique entre montaguiens et chomskyens dans lequel la Grammaire de Montague a été présentée, par lui-même et par ceux qui ont eu la patience de la décortiquer et de l'analyser²².

En schématisant la polémique, elle peut se résumer ainsi :

Montague cherche à montrer que les langues naturelles ne sont pas différentes des langues artificielles de la logique alors que Chomsky, à l'opposé, cherche à montrer leur singularité.

Ainsi présentés, les deux objectifs semblent irréductiblement en contradiction. Mais on peut négocier, grâce à l'homomorphisme-u, un cadre où chacun puisse être considéré comme complémentaire de l'autre.

Si l'on a la trilogie :

- i Alg-L où algèbre décrivant syntaxiquement les expressions d'un langage L
- ii Alg-U, où algèbre décrivant l'univers de dénotation
- iii Homomorphisme-u, associant Alg-L et Alg-U.

nous avons une situation où (i) va varier en fonction des variations de (ii) et ceci si l'on souhaite respecter (iii). Supposons que l'on ait deux langages L1 et L2. Si l'on fixe (ii) et que l'on se donne (iii) comme exigence aussi bien pour <Alg-L1, Alg-U>, que pour <Alg-L2, Alg-U>, les différences entre L1 et L2 peuvent s'exprimer en caractérisant les différences entre Alg-L1 et Alg-L2, les deux étant requises pour satisfaire (iii) sur un même objet ou sur une classe d'objets. Dans le cas qui nous occupe, et en première approximation, si ayant fixé (ii) on avait Alg-LN et Alg-LL (Langages Logiques), on

²² Cf. UG p. 223 note 2, [Partee 97 Section 2], [Halvorsen & Ladusaw 79 p. 185-189], [Dowty & al 81 chap 8].

pourrait caractériser empiriquement LN par rapport à LL en spécifiant, p. ex., ce dont a besoin en Alg-LN et ce dont on n'a pas besoin en Alg-LL pour obtenir (iii).

Montague, avec son Idée 2 citée dans la Section 2 (« L'intérêt principal de la syntaxe est de conduire à la sémantique »), a fait le choix de ne pas utiliser des jugements portant sur la syntaxe comme objets susceptibles de tester l'adéquation externe des descriptions linguistiques. Il le dit moyennant le texte suivant, déjà cité, et que nous reproduisons ici :

It is to be expected [...] that the aim of syntax could be realized in many different ways, only some of which would provide a suitable basis for semantics. [...] I fail to see any great interest in syntax except as a preliminary to semantics [UG, p. 223 note 2].

À partir de cette Idée 2 et en appliquant le schéma que l'on vient de présenter pour situer langues naturelles et langages logiques, on peut dire que la syntaxe des langages logiques conduit à la sémantique comme une avenue spacieuse, bien éclairée, sans croisements, avec une signalisation non ambiguë conduirait à un monument principal auquel il s'agit d'accéder, alors que la syntaxe des langues naturelles conduirait à une esquisse de bâtiment dont les contours ne sont pas clairs, comme le ferait un chemin ardu, qui souvent se perd dans le maquis, chemin non éclairé et sans signalisation ou, pire encore, avec une signalisation trompeuse car souvent ambiguë. L'avenue royale et le chemin ardu conduisent tous les deux à des lieux qui se ressemblent, mais l'une le fait sans difficulté majeure, alors que l'autre est plein d'obstacles à surmonter qui exigent l'utilisation d'outils compliqués. Les difficultés du chemin ardu représentent ainsi la naturalité de la langue naturelle.

Le schéma que l'on vient de présenter n'a rien de révolutionnaire. Ce n'est au fond qu'une application quelque peu étendue des définitions par *genus et differentia*. Les algèbres et l'homomorphisme- \cup nous donnent le *genus* alors que la comparaison des algèbres nous donne la *differentia*.

Ce cadre classique concernant les définitions doit être complété, pour le rendre opérationnel sur le plan empirique, par deux choix concernant (i) l'utilisation empirique des définitions et (ii) l'adoption d'une Alg-U. Les deux choix sont considérés comme indispensables pour pouvoir aboutir à une caractérisation empirique des LN en relation aux LL dans le cadre proposé ci-dessus.

Une définition quelconque peut être utilisée pour caractériser de manière souveraine des objets et les manipuler par la suite. Ces objets n'ont que les caractéristiques que la définition leur a assignées. Tel objet s'adapte ou non à telle définition; celle-ci n'est pas pour ainsi dire mise en cause par les objets ainsi délimités. Mais si on connaît observationnellement une classe d'objets dans l'empirie, une définition peut être observationnellement adéquate ou non selon que tous et seulement les objets que l'on considère comme relevant d'une certaine classe sont couverts par la définition proposée. Si tel n'était pas le cas, on dirait que la définition a échoué et qu'il faut la modifier.

Pour que le cadre proposé plus haut fonctionne empiriquement, il faut l'utiliser en supposant des définitions observationnelles adéquates à ce que sont les langues naturelles et à ce que sont les langages logiques ou, en disant la chose autrement, on doit supposer que l'on peut reconnaître dans l'empirie ce qu'est une langue naturelle et ce qu'est un langage logique²³.

On part de l'idée que les langages logiques ne sont pas empiriquement vrais ou faux par rapport à l'empirie, mais formellement corrects ou non, ce qui est une adéquation différente de l'adéquation descriptive ou explicative par rapport à l'empirie. Pour juger un langage logique on s'interrogera sur sa complétude, sa cohérence et aussi sur son pouvoir expressif lui permettant d'être utilisé pour exprimer tel type de déduction dans tel système d'hypothèses. De telle manière que se pose la question du choix de l'Alg-U avec laquelle les algèbres Alg-LN et Alg-LL doivent être associées par l'homomorphisme-u.

Une des idées-clés du cadre méthodologique dans lequel 5P se situe (cf. Section 8, suivante) est que n'existe pas LA sémantique (ni d'ailleurs LA morphologie, ou LA syntaxe) des langues naturelles. Celle que l'on voudra bien expliciter va dépendre du réel-observé, lui-même conditionné par les choix effectués sur la manière d'observer le réel. Cette position concourt aussi à caractériser la problématique du choix de l'Alg-U, car c'est ce choix qui va déterminer les exigences d'adéquation que l'on voudra se donner pour *la* (en minuscule) sémantique qu'on veut obtenir.

Pour régler cette question des ambitions à fixer à la sémantique, nous avons adopté la position suivante. Nous supposons que nous restons dans le cadre des observations sur les langues naturelles qui peuvent être exprimées par des formules LI et nous supposons que ces observations sont effectivement exprimables ainsi. Par ailleurs, nous nous situons dans le cadre suivant (rappel : Alg-LL et Alg-LN sont les algèbres décrivant la syntaxe de respectivement LL et LN, Alg-U est l'algèbre associée à l'univers de dénotation, AD^{LI} est l'algèbre dérivée de LI par utilisation de la classe K des opération polynomiales) :

- i Alg-LL - homomorphisme-u -> Alg-U

- ii Alg-LN - homomorphisme-u -> Alg-LI
 - ↓
 - Op-K
 - ↓
 - AD^{LI} - homomorphisme-u -> Alg-U
 - + Postulats de signification

²³ Remarquons que toutes les discussions sur cette question font les mêmes suppositions, mais implicitement. On ajoute que nous visons des langages logiques effectivement utilisés en tant que tels pour en éviter d'autres que l'on pourrait inventer pour le besoin de la discussion.

Nous supposons que $Op-K(Alg-LI) = AD^{LI} \neq Identité(Alg-LI)$. La ligne (ii) du tableau n'est autre chose que la présentation graphique de la notion d'interprétation induite de UG avec en plus les postulats de signification.

Dans ce cadre nous pouvons considérer tout langage logique; à noter donc qu'Alg-LI est d'une part le point de départ de Op-K dans la ligne (ii) et aussi un cas particulier de Alg-LL dans la ligne (i). Le cadre devrait permettre de situer les formulations I à IX qui suivent ci-après, et qui s'inscrivent, toutes, dans le schéma sous-jacent suivant.

- Étant donné tout LL et son univers d'interprétation, l'Alg-LL et l'Alg-U qui leur sont respectivement associées, par rapport à Y on observe X dans l' Alg-LL.
- Étant donné toute (ou toute°, sur la différence de *toute* et *toute°*, cf. ci-dessous) LN, si Alg-LN est observationnellement adéquate et qu'elle doit être associée par un homomorphisme-u à LI, par rapport à Y on observe Z dans l'Alg-LN,
où
 - Y est un type de caractéristiques avec lesquelles on peut décrire les « langages » (qu'ils soient « ambigus » ou non) avec leurs univers d'interprétation et/ou leurs algèbres associées
 - X et Z sont de type Y et en contradiction.

Comme ce schéma est supposé sous-jacent aux formulations I à IX qui suivent, celles-ci seront exprimées de manière plus simple et plus intuitivement accessible. On dira simplement :

- α tout LL ... β
- α toute (toute°) LN ... δ

Dans les formulations I à IX on utilise *tout* et *toute* avec leur interprétation d'universel. Par ailleurs, on utilise *toute° LN* dans les formulations I, II, IV, VII, VIII, IX ; cette expression est à interpréter ainsi : dans toute LN on a au moins une des caractéristiques explicitées dans les formulations I, II, IV, VII, VIII, IX.

On suppose que le type de caractéristique Y utilisée dans la comparaison de LL avec LN dans chaque formulation ressort clairement. Ainsi, p. ex, la formulation I est un raccourci pour dire que dans le « langage non ambigu » associé à tout LL il n'y a pas de Op qui permute, alors qu'une opération de ce type est nécessaire dans une Alg-LN.

Les formulations I à IX suivent.

- I
- Dans tout LL il n'existe pas d'opération de permutation.
 - Dans toute° LN une opération de permutation est nécessaire.

Dans la discussion des Sections précédentes, cette formulation s'instancie dans les exemples portant sur les modifieurs, que ceux-ci soient des syntagmes prépositionnels (comme *in the park*) ou des adverbes, de phrase ou verbaux. La situation générale relève des faits suivants. En LI il n'y a pas d'opération de permutation et le foncteur précède toujours l'argument. Or en LN du fait de l'ordre libre ou semi-libre des constituants, ce qui va être traduit en tant que foncteur précède parfois et suit parfois son argument. Les exemples utilisés pourraient s'étendre, p. ex. aux interrogatives.

- II
- Dans tout LL il n'existe pas d'opération d'effacement.
 - Dans toute^o LN une opération d'effacement est nécessaire.

Les opérations d'effacement, comme celles de permutation, sont parmi celles qui augmentent le pouvoir expressif d'un langage au détriment de sa computabilité; en ce qui concerne LL, cette caractéristique pourrait être déduite ou considérée comme une conséquence de V.

Dans les LN, cette caractéristique vient, parmi d'autres, de l'existence d'expressions élémentaires graphiques catégoriellement ambiguës. On sait que, selon les restrictions actuelles, on ne peut pas les traduire tout en restant dans un système observationnellement adéquat (cf. l'Annexe). Pour les faire changer de catégorie, il faut un *vacuum*, qu'il faut effacer. Au-delà de ce cas, on efface aussi pour exprimer l'ambiguïté de portée des quantificateurs.

- III
- Pour tout LL R est l'identité.
 - Pour toute LN R n'est pas l'identité.

Cette formulation utilise les notions de « langage non ambigu » et de « langage ambigu » obtenu celui-ci à partir du premier au moyen de l'opération R. Le constat de III vient de la construction artificielle – dans le sens d'explicitement réfléchi – de la syntaxe des LL : la syntaxe de ces langages a été systématiquement définie en évitant les ambiguïtés. Aucun processus de construction réfléchi n'a existé pour les LN. À chaque moment de leur observation synchronique, une LN est le résultat d'un processus qui, de manière quelque peu machiavélique, aboutit toujours à produire des ambiguïtés, celles-ci étant parmi les fardeaux les plus lourds à traiter dans une Alg-LN. Autrement dit, les LL sont nés non ambigus, de telle manière que R est innécessaire, alors que les LN, dans leur évolution constante, perpétuent les ambiguïtés.

- IV
- Pour tout LL, si on utilise les parenthèses pour désambiguïser, alors ce sont des parenthèses nues (elles ne sont pas distinguées selon l'opération utilisée).

- Pour toute° LN, si on utilise les parenthèses pour désambiguïser, alors ce sont des parenthèses indexées.

Dans les LL, l'opération utilisée est celle de la concaténation et les parenthèses vont être utilisées pour désambiguïser les concaténations effectuées sur une même suite de symboles. Soit, p. ex, (i) : l'utilisation des parenthèses va nous permettre de distinguer (ii) de (iii) :

- i $A \wedge B \cup C$
- ii $(A \wedge B) \cup C$
- iii $A \wedge (B \cup C)$

Dans les LN, puisqu'on a I, II et qu'en plus on utilise la concaténation, une suite quelconque va pouvoir être engendrée de plusieurs manières différentes. Ainsi la suite (i) va, p. ex., pouvoir être engendrée par la permutation de A et B dans une suite (ii), par l'effacement de D d'une suite (iii), ou par la concaténation de la suite A en iv-a avec la suite BC en iv-b :

- i A B C
- ii B A C
- iii A B C D
- iv-a A iv-b B C

LL et LN utilisent donc des parenthèses aussi bien l'une que l'autre, mais l'exigence d'expressivité pour l'un est implacablement plus forte que pour l'autre.

- V
- Tout LL est computable
 - Toute LN n'est pas computable

Ce constat est à mettre en relation avec I et II ci-dessous et par ailleurs, avec les manipulations de pSLF.

- VI
- Dans tout LL, l'assignation catégorielle des expressions graphiques élémentaires se fait par leurs caractéristiques graphiques, éventuellement associées aux nombres naturels.
 - Dans toute LN, l'assignation catégorielle des expressions graphiques élémentaires ne peut pas se faire par leurs caractéristiques graphiques, éventuellement associées aux nombres naturels.
- VII
- Dans tout LL il n'y a pas des expressions graphiques élémentaires catégoriellement ambiguës.
 - Dans toute° LN il y a des expressions élémentaires catégoriellement ambiguës.

Les constats des formulations VI et VII se complètent mutuellement, mais l'un n'est pas déductible de l'autre. Dans la présentation de la syntaxe des LL, la norme est de caractériser par la typographie les entités sur lesquelles on va travailler; si l'on a besoin des ensembles finis dénombrables, on indexe les entités par les nombres naturels (comme p. ex. les *he_i* en pSLF). De telle manière que, d'un simple regard pour ainsi dire, si on connaît les conventions, on sait avec quelle catégorie d'entité on travaille. Ceci n'est pas le cas dans LN. D'une part on a ce qu'on a caractérisé comme la double articulation pour différencier les signifiants des entités lexicales. D'autre part, il n'existe la possibilité, dans aucune langue, de classer les signifiants doublement articulés en fonction des unités de deuxième articulation qui les composent (lettre ou phonèmes) de manière à pouvoir obtenir les catégories syntaxiques adéquates. De telle manière qu'en se référant, p. ex. aux types de syllabe, aux lettres utilisées, on ne peut pas catégoriser les entités graphiques élémentaires des LN et les classer en noms, verbes, etc., c'est-à-dire dans toutes les catégories dont on a besoin pour obtenir une algèbre adéquate²⁴. Par ailleurs le constat dans VII introduit une caractéristique supplémentaire par rapport à celui de III : non seulement pour les LN R n'est pas l'identité mais, en plus, il y a une autre source d'ambiguïté qui est l'assignation ambiguë des expressions élémentaires (cf. l'Annexe).

- VIII • Pour tout LL il n'est pas nécessaire d'utiliser les postulats de signification.
 • Pour toute° LN il faut utiliser les postulats de signification.

Les postulats de signification sont requis pour LN du fait de l'utilisation de *s* dans les types de LI, et aussi du fait que des entités (notamment verbes) qui ont un même comportement syntaxique – et donc qu'il ne faut pas distinguer en syntaxe, comme p. ex. *manger* et *chercher* – exigent des interprétations différentes par rapport à l'extension/intension. Montague adopte la stratégie de la pire hypothèse. Cela veut dire qu'on surgénère des lectures qu'il faudra bien effacer quelque part : ce sont les postulats de signification qui s'en chargent.

- IX • Pour tout LL, l'interprétation peut se faire sur l'Alg-U et non sur une algèbre différente de l'Alg-U dérivée par des opérations polynomiales de l'Alg-U.
 • Pour toute° LN, la traduction sur LI et l'interprétation induite correspondante doit se faire sur des algèbres dérivées à partir des opérations de LI et de l'Alg-U.

L'exigence de trouver des descriptions observationnellement adéquates implique que toute opération pour ainsi dire simple de LI ou dans l'univers de dénotation ne soit pas associée à

²⁴ Cf. dans [Morrill 94 p. 4] une observation analogue : « A principle difference between artificial languages and natural languages is that while the symbols of the former are usually designed to be unambiguous, those of the latter may [d'après nous, il y en a qui le sont dans toutes les langues naturelles] be ambiguous ». Or, malgré le fait que « The definition by induction of model theoretic semantics is undermined by ambiguity », [Morrill 94 p. 4] conclut à une absence de différence significative entre langues naturelles et langages artificiels : « Nevertheless, we follow Montague in denying that description of natural languages requires a significantly different approach than the description of artificial languages ».

une opération sur les structures syntaxiques requises de la langue naturelle. Ceci oblige à passer, dans le cas des LN, par les algèbres dérivées.

Puisque nous prétendons à l'empiricité, il se pose la question de savoir si les formulations I à IX précédentes peuvent se tester. La réponse est oui : s'agissant de formulations qui se veulent empiriques, on peut les tester mais on ne peut pas les prouver. Cela veut dire qu'à un moment quelconque, toute LN effectivement examinée devrait respecter I à IX, ce qui ne garantit nullement qu'il n'y en ait pas encore une, non examinée à ce moment, qui ne les satisfasse pas ou qui ne les satisfasse pas toutes. Toute galaxie examinée jusqu'à nos jours satisfait une relation entre espèces chimiques la constituant, température et distance par rapport au point du big-bang. L'hypothèse du big-bang est ainsi confirmée. Aujourd'hui on a repéré des millions de galaxies que l'on ne connaissait pas il y a quelques années. Personne ne peut prouver ce que seront les observations sur les galaxies à découvrir dans les années qui viennent. Mais ainsi sont les choses dans les sciences empiriques. Comme indication que les formules I à IX ne sont pas erronées, on peut faire le constat suivant : on ne connaît aucun développement, amélioration, modification ou adaptation de la Grammaire de Montague avec incidence sur l'algèbre de la syntaxe utilisée pour décrire les langues naturelles ayant réussi à produire une description de celles-ci incorporant les opérations syntaxiques formelles et computables – et non pas des manipulations du type de celles employées dans pSLF – d'une algèbre susceptible d'être associée par un homomorphisme- u à LI tel que celui défini dans UG et ceci dans un domaine observationnel incorporant le domaine visé dans pSLF²⁵.

Nous soulignons deux points :

- Il s'agit de l'homomorphisme- u tel qu'il est précisément formulé dans UG (et non dans PTQ); il ne s'agit donc pas d'une interprétation quelconque donnée à la compositionnalité en tant que notion fonctionnant comme une auberge espagnole d'autrefois.
- L'expression *toute LN* est plus aisément réfutable que l'expression *toute^oLN*.

Le recours à *toute^oLN* peut se justifier intuitivement en observant la ligne (ii) du tableau ci-dessus : pour pratiquer l'homomorphisme- u entre la langue naturelle et l'univers d'interprétation, le système d'ensemble se donne trois zones (en plus de la relation R irrestrie) pour évacuer les difficultés : l'Alg-LN elle-même, les Op-K (qui ne sont limitées par aucune contrainte) et les postulats de signification. Autrement dit, ce que l'expression *toute^oLN* exprime est que, pour pallier une difficulté de description d'une partie de la syntaxe des langues naturelles, le système est suffisamment peu contraint pour permettre trois zones d'intervention, sans restrictions ni sur ce qui revient à chaque zone ni sur le pouvoir expressif de chaque zone. D'où le besoin d'avoir recours à des alternatives,

²⁵ Cette remarque vaut, p. ex., pour [Dowty 79], [Dowty 82], [Bach 79], [Partee 79a], [Partee 79b], [Cooper 83]. Cf. ci-dessous Note complémentaire 1 et, autour du traitement des « logical words », dans la Section suivante, une autre différence entre langages logiques et langues naturelles.

ce qu'exprime la notation *toute*^o*LN* : on peut, p. ex. se donner des opérations polynomiales plus compliquées dans IX et s'éviter des permutations dans I.

8 5P et compositionnalité

Dans les Sections 1 à 6, nous avons revisité la Grammaire de Montague à partir de deux points de vue, celui de l'empiricité et celui de la compositionnalité, et ceci aussi bien dans les déclarations générales de UG que dans la pratique effective de PTQ. Nous avons été ainsi conduit à présenter des éléments permettant de réévaluer la Grammaire de Montague. Dans la Section 7 nous avons essayé de renverser le statut méthodologique de la compositionnalité pour en faire une utilisation empirique. À partir d'un constat observationnel que l'on peut accepter – on sait que tels objets sont des langues naturelles et que tels autres objets sont des langages artificiels des logiciens –, on a essayé de caractériser les différences entre les deux, et ceci, en particulier, en précisant leur différence au niveau de la syntaxe si l'on veut faire en sorte que les langues naturelles soient associées moyennant un homomorphisme-*u* à LI. Il nous reste dans cette dernière Section à proposer les choix sur la compositionnalité dans la spécification de la sémantique de 5P.

Si 5P se présente sous la rubrique quelque peu prétentieuse de *paradigme* (notion empruntée à [Kuhn 83]), c'est moins – à la différence de la notion originale – du fait que l'on se propose de regarder dans l'objet d'étude des choses différentes de celles que l'on regarde d'habitude (la bonne formation, les situations de paraphrase, d'ambiguïté, le calcul des conséquences, etc.) que parce qu'on souhaiterait traiter notre objet d'étude – les langues naturelles – avec les exigences communément admises comme socle minimum par les sciences empiriques. Les exigences méthodologiques de 5P devraient pouvoir se déduire, toutes, de l'idée simple suivante : la linguistique est une science empirique, la linguistique telle qu'elle pourrait se pratiquer, plutôt que la linguistique telle qu'on la pratique effectivement.

De cette idée directrice découlent, nous semble-t-il, les idées suivantes.

- i L'observation systématique et l'adéquation par rapport à l'adéquation prime par rapport à tout critère ou métacritère ou formule évaluative que l'on peut se donner sur la formulation des hypothèses descriptives et/ou explicatives, sous couvert que celles-ci soient formulées de manière explicite, calculable et cohérente.
- ii L'observation n'est pas le résultat d'une évidence mais d'une construction faite à partir de points de vue dûment explicités.
- iii Dans un système avec une multiplicité d'hypothèses qui concourent à l'obtention d'un résultat final, multiplicité telle que la décision sur une hypothèse d'un certain type entraîne quasiment fatalement des répercussions sur les autres, il est strictement vain

d'essayer de justifier chaque hypothèse ou chaque type d'hypothèse isolément. C'est l'interaction de l'ensemble des hypothèses qui provoque ce résultat final qui va être confronté (testé, contrasté) par rapport au réel-observé; nous ne sommes pas dans un cas de figure où chaque segment des hypothèses peut être confronté et justifié par rapport à un segment du réel-observé.

À partir des Sections précédentes et du rappel des ambitions méthodologiques de 5P, il est clair que nous ne nous sentons pas engagé dans la spécification de la sémantique de 5P par le commandement méthodologique *homomorphisme syntaxe-sémantique tu construiras*. Et ceci pour les deux grandes raisons que l'on peut résumer ainsi :

- Si l'on veut utiliser la compositionnalité^{hu} dans le cadre d'une science empirique – comme esquissé dans la Section 7 – il faut la subordonner à l'adéquation sur l'observation, afin de pouvoir évaluer ce que cela coûte de la respecter si l'on veut couvrir un champ donné d'observations, ce coût étant mesuré par l'augmentation du pouvoir expressif requis pour décrire la syntaxe et/ou la dénotation; c'est cette augmentation dans le « coût » des entités utilisées qui va nous donner une caractérisation empirique de la différence entre langue naturelle et langages artificiels.
- Malgré des affirmations dans le sens contraire, en particulier dans [Janssen 97], on ne voit strictement rien de mauvais ou de dangereux qui ait été évité par la compositionnalité^{hu}. En revanche, on a pu observer les malentendus que la notion générale et de type « auberge espagnole d'autrefois » de compositionnalité a introduits.

Ce qui précède peut justifier la désobéissance au commandement de compositionnalité à la Montague et/ou à la Janssen dans la spécification de la sémantique dans 5P, désobéissance fondée dans l'objectif d'aborder la linguistique comme science empirique. Nous esquissons par la suite comment il est possible de procéder.

Dans le schéma classique induction-déduction des sciences empiriques, il existe une interaction constante entre observation, formulation des hypothèses, calcul des déductions à partir des hypothèses et vérification des déductions en les confrontant à des observations, celles-ci étant effectuées soit avant soit après la formulation des hypothèses.

Les cinq *p* du Paradigme 5P sont là pour concrétiser ce schéma général en relation à l'étude des langues humaines. Il sont déclinés par la suite²⁶.

- *P1* ou *P* de Protocoles, un Protocole étant la représentation d'une donnée à laquelle aboutit un Observateur dûment modélisé ; les *P1* (Protocoles) sont donc le réel-observé de §3.1.

²⁶ Dans [Bès 99] on trouvera la présentation du Paradigme 5P avec une application détaillée à la phrase verbale noyau du français. [Dans Bès & Hagège 01], travail qui devrait être bientôt installé dans la page Web du GRIL, la présentation, formelle de 5P, et notamment des *P2* et du système CAT, est nettement améliorée; on y trouvera aussi les antécédents de 5P et les travaux qui l'ont utilisé. Dans ce même [Bès 99] ont été présentés les choix de fond sur le calcul de la sémantique, notamment l'utilisation du fléchage en entrée des Fonctions sémantiques, l'option d'une compositionnalité du type compositionnalité^{strad} et l'objectif de pouvoir aboutir en sortie des Fonctions sémantiques à une pluralité de formules logiques.

- *P2 ou P* de Propriétés, une Propriété pouvant s'identifier formellement à un axiome. Un ensemble fini de Propriétés spécifie en intension un ensemble de suites d'une langue donnée. On appelle *modèles* ces suites. Un *modèle* est ainsi une suite qui satisfait un ensemble de Propriétés.
- *P3 ou P* de Projections, une Projection étant une généralisation sur les Propriétés ou sur un sous-ensemble de Propriétés d'une langue naturelle.
- *P4 ou P* de Principes, un Principe est une contrainte sur les Projections valable pour toutes les langues, dans les meilleurs des cas, ou pour un groupe important des langues.
- *P5 ou P* de Processus, un Processus étant une procédure effective implantable ou implantée en machine permettant de traiter des suites LN.

Supposons un ensemble de Propriétés (P2) et un ensemble non vide de modèles qui satisfont cet ensemble de Propriétés. On utilise des identifiants qui vont permettre de référer à des ensembles de Propriétés et aux modèles qui les satisfont. Par exemple, l'identifiant *N* sera utilisé dans la notation *Pr-N* pour noter l'ensemble de Propriétés qui vont spécifier les modèles qui sont des syntagmes nominaux, la notation *M-N* pour noter l'ensemble des modèles qui satisfont *Pr-N*, et la notation *m-N* pour noter un quelconque élément de *M-N*. Supposons par ailleurs l'ensemble *FL* (Formes Linguistiques) dont chaque élément – *fl* – est une suite de caractères, c'est à dire une expression graphique nue. Par exemple, on aurait dans *FL* en français *les, si, artiste, changeons, ...*

À partir de *FL* on peut spécifier FL^+ , l'assemblage sans symbole nul de toutes les suites que l'on peut former à partir de *FL*; on note *sfl* un élément de FL^+ . L'Observateur devra associer à un élément de FL^+ des observations représentées, que nous notons *Obr*. Nous obtenons ainsi la paire $\langle sfl, Obr \rangle$. Le réel-observé de §3.1 est un ensemble constitué par des paires $\langle sfl, Obr \rangle$.

On synthétisera ci-dessous les différents types des Propriétés que nous distinguons dans les P2 de 5P. Acceptons que chaque type de Propriétés contribue à enrichir les modèles qu'elle contribue à spécifier et que le symbole *Inf* synthétise toutes les informations associées à une *slf*. Étant donné un ensemble de Propriétés P2, nous aurons un ensemble de $\langle slf, Inf \rangle$, spécifié à partir de P2.

Nous avons ainsi des objets relevant de l'observation, les $\langle slf, Obr \rangle$, et des objets relevant de la description hypothétique, les $\langle slf, Inf \rangle$. Il nous faut ajouter le dernier concept, celui de *traduction observationnelle*, qui va être une mise en relation de *Obr* avec *Inf*, étant donné une paire

$$\langle \langle slf_i, Obr \rangle, \langle slf_j, Inf \rangle \rangle, \text{ avec } F(slf_i) = (slf_j)$$

F étant dans le cas plus simple – que nous supposons ici – l'identité. Les représentations sémantiques relèvent de *Inf*, ce dont elles doivent rendre compte relève de *Obr*. On observe que cela implique d'utiliser *FL* pour aussi bien *slf_i* que *slf_j* ce qui semble assez justifié lorsqu'on travaille avec des textes sur support magnétique, comme c'est notre cas, mais qui ne le serait pas du tout si l'on s'adressait au langage parlé.

Dans les discussions sur la sémantique on justifie tel ou tel type de description sur un énoncé avec des anecdotes sur l'interprétation exigée pour le décrire : c'est ce que [Janssen 97] nomme des « histoires » (anglais « story ») ; cf. ci-dessous dans cette même Section un exemple. Ces « histoires » peuvent aussi être dites sous forme de gloses ; cf. une utilisation ponctuelle dans l'Annexe ci-près, §2.1 in fine. Mais on ne dispose pas dans la situation actuelle d'un vrai langage de spécification pour représenter systématiquement les aspects sémantiques observés. L'observation sur forme représentation de ce dont une représentation sémantique doit rendre compte est, certes, comme il est dit par Cooper (cf. ci-après Note N° 2), un problème subtil et compliqué, et il l'est tellement – et pas seulement pratique²⁷ – qu'en général on fait comme s'il n'existait pas.

Nous avons pris position par rapport aux trois idées basiques de Montague présentées ci-avant §2. L'idée 1 (« Il n'y a pas de différences théoriques fondamentales entre les langages naturels et les langages artificiels des logiciens ... ») a été considérée comme non empirique; cf. dans le premier alinéa de §7.2 un résumé des motivations pour dénier son empiricité, et, pour les raisons ci-dessus exposées, nous ne nous plions pas au commandement méthodologique de [Janssen 97]. De cela découle une conséquence non triviale, à savoir la nécessité d'accepter une représentation intermédiaire entre les expressions graphiques et l'univers de dénotation.

En effet, nous savons que la notion-clef d'interprétation induite de Montague est celle qui établit la dispensabilité de principe de LI (cf. ci-avant §3.2.1 et en particulier le point (xiii) de UG) : grâce à cela, le système en cinq éléments de UG n'est pas une nécessité de principe mais plutôt une commodité d'écriture, ce qui permet de conserver comme seuls éléments en principe nécessaires les trois de EFL (cf. § 1.1). Le raisonnement général est le suivant :

- Étant donné une description syntaxique d'une suite *s* de LN obtenue par un langage « non ambigu » et les exigences de traduction sur LI, l'interprétation induite de *s* via LI va à aboutir dans l'univers de dénotation à la même interprétation que si on procédait directement de LN vers cet univers.

C'est ainsi que [Dowty & al. 81] prend soin d'explicitier qu'une form-LI à laquelle on aboutit par la traduction ne constitue pas un « vrai » niveau de représentation d'une suite-LN (c'est nous qui soulignons en gras) :

It will probably be tempting for the reader approaching Montague Grammar for the first time to think of the IL translation of an English sentence as its “semantic representation” or its “logical form”. Indeed, it may be helpful at first to make this association in order to get a better intuitive grasp of the system, but one must ultimately come to realize that this translation plays quite a different and, in many ways, less significant role in Montague’s program than does the concept of logical form within certain linguistic theories. **For Montague, the translation is in fact a**

²⁷ Les gloses sont exprimées dans LN, vraisemblablement par un sous-ensemble de structures syntaxiques et d'entrées lexicales d'une LN, utilisées en tant que langage de spécification ... qui ont à leur tour des rsem qui restent en dehors de ce qu'elles spécifient. On peut retrograder aussi loin que l'on voudra, il restera toujours un sous-ensemble de la LN non formalisée : voilà encore un point de différence – et non trivial – avec les langages artificiels des logiciens.

completely dispensable part of the theory. The “real” semantic interpretation of an English sentence is simply its model-theoretic interpretation, however arrived at, and nothing else (p. 181).

Translating English into Intensional Logic was therefore not essential to interpreting the English phrases we generated ; it was simply a convenient immediate step in assigning them meanings. This step could have been eliminated had we been chosen to describe the interpretation of English directly, rather than indirectly by stipulating that it be the one induced under our translation into Intensional Logic by the interpretation given to that language. This point is important, because anyone who does not appreciate it may misunderstand the role of Intensional Logic in applications of Montague’s descriptive framework to natural languages.

Some writers have suggested that Intensional Logic plays a role in “Montag Grammar” like that of Logical Form in Generative Semantics or in Chomsky’s current conception of grammar [1977]. This is definitely not the case. **The idea is fundamentally mistaken in that translation into Intensional Logic is, as we have just seen, not in any way essential to semantic interpretation in Montague’s framework** – that step of PTQ is entirely eliminable without altering anything crucial about the interpretation assigned each phrase. It is further mistaken in that the particular translations assigned play no role in determining well-formedness of the syntactic structures with which they are associated, as Logical Form may do in Chomsky [1980]. And it is also mistaken in its implication that the entailments licensed by a sentence of English are formalized by the translations of that the sentence into Intensional logic and the inference rules of that logic (p. 263-264).

Ce raisonnement peut être compris dans sa variante plus forte : si l'on veut interpréter les expressions LN dans un univers de dénotation, la dispensabilité de principe d'une représentation intermédiaire ne peut s'obtenir que si la traduction de la syntaxe des expressions LN sur la représentation intermédiaire satisfait les exigences de la compositionnalité^{hu}. Sans qu'il soit prouvé formellement, on peut l'accepter comme une conjecture raisonnable, de telle manière que 5P aurait la nécessité de principe d'avoir recours à une structure intermédiaire.

Quoiqu'il en soit au niveau des raisonnements formels, on n'essaye pas en 5P de se dispenser d'une représentation sémantique (désormais *rsem*) entre la syntaxe des expressions LN et l'univers de dénnotations dans lequel elles doivent être interprétées.

Quelle *rsem*?

Pour poser le problème de la naturalité de la langue naturelle dans §7.2, le choix a été expéditif : une *rsem* est une form-LI. Or cette question nous ramène à l'Idée 3 de Montague (cf. §2) selon laquelle « l'objectif principal de la sémantique est de caractériser les notions d'énoncé vrai et de conséquence ». Le choix des form-LI en tant que *rsem* s'inscrit dans cette Idée 3 qui est elle-même le résultat d'un choix plus général pour observer le Réel, un choix d'après nous et non une nécessité, les justificatifs de « bon sens » pour étayer sa nécessité étant très faibles (cf. §3.1). D'autres choix que ceux qui relèvent des valeurs de vérité et des déductions afférentes sont possibles et intéressants, mais nous ne les aborderons pas ici²⁸.

²⁸ Le traitement de la paraphrase est centrale pour le modèle Sens-texte (cf. parmi d'autres [Mel'čuk & Pertsov 87], [Mel'čuk 87], et on voit mal pourquoi dénier aux Observations sur la paraphrase le statut des données susceptibles de

Dans le cadre de l'Idée 3, LI est un choix possible et justifié par sa capacité expressive, mais ce n'est pas le seul possible. La logique de premier ordre, les DRT à la Kamp sont, parmi d'autres, des possibilités à envisager. Il n'existe pas selon cette manière de voir les choses, LA sémantique de la LN. On a une pluralité de possibilités dictées au départ par la manière d'observer et par ce qu'on veut faire à la sortie avec les rsem obtenues.

Cette relativisation des rsem ne manque pas de se répercuter sur l'Idée 2 de Montague, selon laquelle « l'intérêt principal de la syntaxe est de conduire à la sémantique », idée à laquelle nous avons adhéré (cf. §3.1). Si nous pensons par ailleurs que les intuitions exprimées par les usagers sont la voie nécessaire ou principale que l'observation doit suivre (cf. Note 2 ci-après), dès qu'on essaye de pratiquer les intuitions sur les objets syntaxiques associés aux expressions décrites, elles deviennent non opérationnelles car non intersubjectives du fait de leur dépendance étroite avec le « modèle » de grammaire à partir duquel ces structures sont construites.

Mais on perçoit la difficulté méthodologique que pose l'adoption de l'Idée 2 : puisqu'on admet une pluralité en principe des sémantiques et que l'intérêt de la syntaxe est de conduire à chacune d'entre elles, est-ce que cela ne revient pas à admettre qu'il y a en principe une pluralité de syntaxes ? Bien entendu, il s'agirait d'une pluralité de syntaxes qui viendrait non des descriptions différentes conduisant à la même sémantique, mais du fait d'être différentes parce que conduisant à des sémantiques différentes. Est-ce qu'il n'y aurait pas ainsi la syntaxe pour LI, la syntaxe pour DRT, la syntaxe pour la logique de premier ordre, etc. ?

La question est plutôt irréaliste dans la situation actuelle, car il n'existe aucune description permettant d'obtenir les rsem sous forme de formules d'une logique quelconque et permettant de couvrir les énoncés des textes effectifs. On sait traiter dans l'une ou l'autre logique des exemples d'école et non des textes effectivement utilisés en tant que tels, si bien qu'aboutir à une seule syntaxe dans une seule logique, permettant d'élargir la couverture de ce que l'on peut représenter, sera déjà très bien.

Mais la question de la nécessité de principe d'une pluralité de syntaxes est néanmoins bien réelle. Pour y apporter des éléments de réponse nous devons présenter avec un peu plus de détails – mais toujours de manière très succincte – l'outillage de 5P pour les descriptions au niveau de P2 (Propriétés).

On sait que la traduction observationnelle va mettre en relation les deux paires suivantes

$\langle\langle \text{slf}_i, \text{Obr} \rangle, \langle \text{slf}_j, \text{Inf} \rangle\rangle$

tester des hypothèses sur la sémantique. Mais la paraphrase et/ou l'ambiguïté ne sont pas les seules questions à pouvoir fournir ce type d'observations, qui ne relèvent pas forcément et directement de la vérité et fausseté des énoncés, bien qu'à un moment donné elles puissent y être intégrées. Ainsi, p. ex., tout ce qui est dit sur l'anaphore en PTQ suppose une identité référentielles des termes anaphoriques. Or sous l'étiquette générale d'anaphore se cachent beaucoup de phénomènes qui relèvent parfois du référent, et parfois du « sens », phénomènes qu'on a réussi largement à décrire et dont l'observation commence à être bien opérationnalisée (cf. [Trouilleux & al 00a], [Trouilleux & al 00b]).

que sfl_i et sfl_j sont des éléments de FL^+ , spécifiée à partir de FL . Les éléments de FL sont les expressions élémentaires de Montague. Un système CAT constitué essentiellement par un alphabet de $\langle \text{étiquette}, \{\text{valeurs}\} \rangle$ et par des relations d'héritage, va spécifier un ensemble – CM – de catégories maximales (cm), chaque catégorie étant un ensemble d'éléments $\langle \text{étiquette}=\text{valeur} \rangle$. CAT détermine un système de subsomption entre catégories. Une fl est associée à une ou plusieurs cm .

Les Propriétés (P2) sont exprimées sur les catégories, maximales ou non; très généralement non maximales, pour profiter de l'expressivité de la subsomption. Un modèle – au sens de 5P – incorpore toujours une liste de catégories; nommons-la *suite-modèle*.

Trois types de Propriétés (P2) sont distinguées :

- Propriétés d'existence, qui spécifient les ensembles de symboles qui peuvent être utilisés dans les suites-modèle ;
- Propriétés de linéarité, qui spécifient les relations d'ordre qui doivent être associées aux ensembles spécifiés par les Propriétés d'existence ;
- Propriétés de fléchage, qui spécifient des ensembles de paires de fléchage.

On essaye dans 5P de pousser la factorisation et la modularisation des informations au maximum, de telle manière que selon l'utilisation de différents types d'information on peut avoir des descriptions plus ou moins sophistiquées.

On reconnaît ainsi trois types de modèles :

- Modèles basiques ;
- Modèles fléchés ;
- Modèles sémantiquement décrits.

En notant en minuscule les catégories, en majuscule l'identifiant du modèle noté en indice (ici N qui correspond au syntagme nominal classique) et en numérotant en indice de chaque catégorie les positions dans la suite-modèle, la sfl du (i) suivant sera associée au modèle basique (ii) et au modèle fléché (iii).

- trois très jolies fleurs très bleues
- $(\text{card}_1 \text{ adv}_2 \text{ adj}_3 \text{ n}_4 \text{ adv}_5 \text{ adj}_6)N$
- $\langle (\text{card}_1 \text{ adv}_2 \text{ adj}_3 \text{ n}_4 \text{ adv}_5 \text{ adj}_6)N, \{ \langle 1, 4 \rangle, \langle 2, 3 \rangle, \langle 3, 4 \rangle, \langle 5, 6 \rangle, \langle 6, 4 \rangle \} \rangle$

En général un modèle fléché a la forme générale suivante :

$\langle (\text{suite-modèle}), \{ \text{ensemble de paires de fléchage} \} \text{Id modèle} \rangle$

Une même suite-modèle peut être associée à plus d'un ensemble de paires de fléchage. Ainsi pour le classique (i) suivant, on aura – de manière très résumée – les paires de fléchage soit dans (ii) soit dans (iii).

- i (Pierre)₁ (a regardé)₂ (les filles)₃ (avec des lunettes)₄
- ii { $\langle 1, 2 \rangle, \langle 3, 2 \rangle, \langle 4, 2 \rangle$ }
- iii { $\langle 1, 2 \rangle, \langle 3, 2 \rangle, \langle 4, 3 \rangle$ }

On obtient ainsi que LF, le système CAT et les Propriétés vont exprimer ce qui correspond à la syntaxe chez Montague, donnant à syntaxe la définition étroite de la théorie générale des langages [Morris, 38]. Moyennant LF, CAT et P2 on doit décrire tout ce qui concerne la bonne formation des expressions graphiques données comme primitives dans LF; autrement dit, LF, CAT et P2 spécifient les « meaningful expressions » de Montague.

Dans tout cela il n'y a rien de sémantique, mais on a la pièce qui va être l'entrée des Fonctions Sémantiques : le graphe constitué par l'ensemble des paires de fléchage. Les Fonctions Sémantiques vont prendre ces graphes en entrée et elles doivent donner les rsem en sortie.

Un Lexique qui, d'une manière ou une autre, est constitué par des entrées lexicales avec, en terminologie saussurienne, un signifiant et un signifié, est conceptuellement un FL associé à un ensemble de Fonctions sémantiques – souvent polymorphiques – associant un signifiant à un ou plusieurs signifiés. Ce type de fonctions est naturellement utilisé dans 5P. Nommons-les FSL (Fonctions sémantiques lexicales). On a ainsi les suites-modèle qui, lorsqu'elles sont entièrement constituées de catégories, vont être associées du côté de la syntaxe, à des *cm*, et par *FSL*, à des rsem. C'est l'équivalent de TR2 dans le système de traduction de Montague (cf. § 4.2.1), à la grande différence près que l'on peut associer une *fl* à plus d'une rsem. Grâce à cette double possibilité – associer une *fl* à plus d'une *cm* et à plus d'une rsem –, on n'a aucun problème pour exprimer l'ambiguïté syntaxique et sémantique des expressions primitives.

Transversalement aux modèles basiques et aux modèles fléchés, on reconnaît les modèles terminés – ainsi nommés lorsque les catégories de la liste-modèle sont des *cm* et qu'elles sont associées à des *fl* – et des modèles terminés lexicalisés, lorsque les catégories sont aussi associées à des rsem. Dans notre exemple précédent

(*card1 adv2 adj3 n4 adv5 adj6*)*N*

card1 sera associé à la *cm* correspondante, à la *lf trois* et à une rsem ; de même pour tous les éléments de la suite-modèle. De cette manière un *n* quelconque dans une paire de fléchage – p. ex. *1* dans la paire <1, 4> de l'exemple – dénote un objet complexe avec l'information sur sa place dans la suite-modèle, la *cm* et la rsem auxquelles il est associé. Le pari essentiel est qu'on a ainsi dans les graphes spécifiés par l'ensemble des paires de fléchage dans les modèles fléchés lexicalement terminés l'information suffisante pour obtenir les rsem.

Ce n'est pas ici le lieu pour justifier les options de 5P, son éventuelle originalité formelle ou sa capacité à obtenir effectivement des rsem sous la forme des formules d'une logique quelconque. Remarquons cependant deux points :

- La capacité expressive formelle requise pour l'obtention de l'ensemble des paires de fléchage et pour l'obtention des rsem permet l'expression des relations contextuelles. De ce fait, on doit pouvoir exprimer la relation entre *lui* et *après* dans (i) qui suit pour obtenir la même rsem que pour (ii) :

- i Jacques lui court après
- ii Jacques court après lui

Les P2 ont donc plus de pouvoir expressif qu'une grammaire hors contexte²⁹.

- Il n'y a pas de processus de dérivation. Les Propriétés sont des analogues des axiomes dans un système axiomatique. Nous supposons qu'il y a basiquement deux types de systèmes axiomatiques : avec et sans règle de substitution. Les Propriétés (P2) sont les analogues des systèmes axiomatiques avec règles de substitution³⁰.

Suite à cette présentation synthétique de l'outillage descriptif de 5P, revenons à la question de la variabilité de principe de la syntaxe, que l'on peut reformuler ainsi : est-ce que le fait d'envisager une pluralité de rsem différentes exige d'avoir autant de syntaxes différentes?

Nous croyons que non, que la variabilité pourra être limitée aux modifications à introduire aux Fonctions sémantiques et non à leur entrée, celle-ci étant la syntaxe exprimée par FL, CAT et P2. Pour justifier intuitivement cette idée, considérons l'énoncé dans (i) qui suit, son rsem plutôt classique en logique de premier ordre dans (ii) et, dans (iii), sa rsem exprimée par une form-LI³¹ à la PTQ.

- i Pierre marche
- ii marcher'(pierre)
- iii $\lambda x [\sim x [\wedge \text{pierre}]] [\wedge \text{marcher}]$

(ii) et (iii) ont l'air très différent – et en quelque sorte, sur le plan de comment on doit comprendre leurs dénnotations, ils le sont. Mais le conditionnement syntaxique qui les détermine peut être « factorisé » et il est le même. En fait, intuitivement, on peut l'exprimer ainsi :

- Phrase verbale avec le sujet comme seul constituant syntaxique jouant le rôle d'argument du verbe, ce sujet étant un nom propre.

5P exprime ce conditionnement utilisent FL, CAT et P2. On retrouve ce conditionnement dans les modèles fléchés terminés, qui seront utilisés dans l'entrée des Fonctions sémantiques; ce sont celles-ci qui doivent donner en sortie les formules telles que celles de (ii) ou (iii).

La présentation de l'outillage descriptif de 5P, et, en particulier, l'inexistence d'une dérivation, permet de revenir au commandement de compositionnalité de [Janssen 97], mais maintenant à partir d'un autre angle. Ce n'est pas seulement que, d'un point de vue méthodologique, l'on ne veut pas pratiquer la compositionnalité^{hu} et/ou la compositionnalité^{janssen} : c'est que tout porte à croire qu'on ne peut pas la pratiquer selon un point de vue formel.

En effet, l'esquisse de caractérisation empirique des langues naturelles en fonction du coût à payer pour satisfaire ce type de compositionnalité et les propositions de [Janssen 97] pour y aboutir coûte qui coûte vont toutes dans une même direction : l'obtention de ce type de

²⁹ Cf Gabriel G. Bès, *5P-Pouvoir expressif des Propriétés (P2)*, GRIL, 1999. Rapport.

³⁰ Cf [Partee & al., 90 chap. 8].

³¹ On suppose la form-LI que l'on obtiendrait à partir de *Pierre (T) marche (IV)* et non de *Pierre (T) heq(T) marche(IV)*.

compositionnalité est garantie possible dans une formulation de la syntaxe qui a le pouvoir expressif des machines de Turing³², ce qui n'est pas le cas des Propriétés (P2).

Mais cela n'est pas tout. [Janssen 97] introduit une autre idée qui, elle, est absolument contradictoire avec ce qui est dans 5P un idée directrice. En effet, pour pouvoir aboutir à la compositionnalité^{hu} désirée, [Janssen 97] ne traduit pas une structure *s* dans une autre, mais il prend en entrée le processus qui permet l'obtention de *s*. Ou autrement dit, [Janssen 97] pose comme exigence pour l'obtention de l'homomorphisme-*u* que le domaine de la fonction de traduction soit le processus représenté par la suite dérivationnelle qui est soigneusement distingué du graphe final que ce processus dérivationnel permet de spécifier. Citons donc le passage clef de [Janssen 97 p. 427], présenté sous forme de « Assumptions », dont nous présentons ici la 8 et la 9, (qui sont par ailleurs quasiment redondantes) :

[Assumption] 8 The meaning of an expression is determined by the way in which it is formed from its parts. The syntactic production process is, therefore, the only input to the process of determining its meaning [...]

[Assumption] 9 The production process is the input for the meaning assignement. Ambiguous expressions must have different derivations: i.e. a derivation with different rules, and/or with different basic expressions³³.

Or l'utilisation du processus dérivationnel réintroduit ce que, dans 5P, on veut soigneusement laisser dehors, à savoir des aspects algorithmiques dans des descriptions qui se veulent purement déclaratives. Les P2 de 5P ne décrivent pas des processus dérivationnels. Elles vont spécifier des modèles, dont les plus achevés sont des modèles avec un fléchage. On n'en récupère pas de dérivations; elles n'existent pas; par conséquent elles ne sont l'entrée de rien. Si l'obtention de l'homomorphisme-*u* en a besoin, la conclusion en est claire : celui-ci ne peut pas être obtenu.

Pour en finir avec cette justification du ne pas vouloir ou du ne pas pouvoir appliquer une compositionnalité^{hu} au calcul de la sémantique dans 5P, remarquons qu'au fond, ce qui est insupportable aux tenants de la compositionnalité^{hu} c'est de ne pas pouvoir associer tel effet de la sémantique à telle opération en syntaxe.

Soit la question de la portée des quantificateurs. Pour rendre compte de l'ambiguïté de (i), la solution dans le cadre de l'homomorphisme-*u* est de l'associer à deux arbres

³² Cf. la démonstration du théorème de [Janssen 97 p. 454-456]. Comme c'est indiqué *ibidem*, on se retrouve avec un langage « which can be generated by the most powerful kinds of grammars 'unrestricted rewriting systems', transformational grammars, Turing machine languages, etc. ».

³³ Sur l'impossibilité non discutée par [Janssen 97] d'utiliser des expressions élémentaires catégoriellement syntaxiquement ambiguës, cf. notre discussion précédente dans la Section 3.2.1 Ob. 3 et l'Annexe; le point en discussion ici n'est pas celui-ci mais l'exigence d'agir sur le processus dérivationnel lui-même et non sur la structure finale que celui-ci spécifie. Notons que [Janssen 97 p. 442] a besoin de ces « Assumptions » pour réfuter [Pelletier 93, 94b], qui a eu le bons sens d'observer que « In order to maintain the Compositionality Principle, theorists resorted to a number of devices which are all more or less unmotivated (except to maintain the Principle) » (propos cités par [Janssen 97 p. 442]).

dérivationnels différents; celui qui correspond à (ii) va mettre l'universel sous la portée de l'existentiel, et vice versa pour celui qui correspond à (iii).

- i Chaque étudiant a lu un livre
- ii (Un livre) (chaque étudiant) a lu (he0)
- iii (Chaque étudiant) a lu (un livre)

On arrive ainsi à caractériser quelle opération en syntaxe correspond à quelle opération en sémantique en évitant d'exprimer l'ambiguïté à partir d'un graphe unique pour (i), ce qui est strictement interdit pour les tenants de la compositionnalité^{hu}.

Mais le problème se complique autrement lorsqu'on sort des cas d'école des exemples classiques et qu'on les étend. Soit p. ex.

- iv Chaque étudiant a lu un seul et même livre
- v Chaque étudiant a lu un livre différent

L'exemple (iv) doit recevoir la même interprétation que (ii) alors que (v) doit être interprété selon (iii) (plus d'autres exigences absentes de (iii)). Or cet exemple illustre ce qui semble être plutôt la situation typique que le cas isolé : les effets en sémantique sont produits par une pluralité de facteurs difficilement isolables les uns des autres. Dans nos exemples, selon la qualification d'identité ou de non-identité des objets dénotés par le syntagme nominal, qualification exprimée par des modificateurs comme *autre*, *même*, *différent* on a des répercussions au niveau de la portée des quantificateurs. Celle-ci est donc bien un résultat d'une pluralité de facteurs et l'on voit mal la possibilité de les isoler pour pouvoir exprimer qu'à telle opération introduisant tel facteur doit correspondre telle portée et pas telle autre : selon la fonction syntaxique du syntagme nominal (sujet, objet, autre type de complément), selon le verbe utilisé, selon l'utilisation des pluriels collectifs ou non, selon la constitution interne du syntagme nominal, etc., on aura vraisemblablement des répercussions au niveau de la portée. Et sans pouvoir démontrer que la compositionnalité^{hu} est impossible, nous croyons qu'il est infiniment plus urgent de bien expliciter ces facteurs et de les traiter formellement dans la construction de la représentation sémantique sans assumer aussi le fardeau de la compositionnalité^{hu}, qui, sur le plan de l'adéquation observationnelle, n'ajoute strictement rien. Remarquons que l'expression de cette pluralité de facteurs est en principe possible en 5P : ils sont tous représentés dans le graphe spécifié par l'ensemble des paires de fléchage.

Le projet sur la sémantique dans 5P peut s'organiser en fonction des trois points suivants :

- observer les énoncés selon un point de vue sémantique et spécifier un langage observationnel avec lequel expliciter les observations ;
- adopter un langage de calcul logique, p. ex. LI, si l'on reste dans le cadre de l'Idée 3 de Montague ;

- traduire³⁴ moyennant les Fonctions sémantiques les modèles associés aux énoncés observés dans les formules du langage logique adopté; cette traduction serait compositionnelle dans ce sens qu'elle doit être explicite et systématique, mais on n'exige pas d'elle qu'elle satisfasse l'homomorphisme-*u*; selon le classement des compositionnalités de la Section 3.2, nous aurons simplement une *compositionnalité*^{strad}.

Nous reprenons par la suite et successivement chacun de ces points.

Pour illustrer la première ligne concernant l'observation nous exploitons deux exemples, l'un sur la coordination et l'autre sur les lectures *de re* et *de dicto*.

Soit l'exemple suivant (i). D'après PTQ, il peut recevoir en pSLF plusieurs structures syntaxiques, parmi lesquelles celles qui sont schématisées dans (ii), (iii) et (iv).

- i Pierre et Nicole et Jacques et Françoise ...
- ii (Pierre et Nicole) et (Jacques) et (Françoise)
- iii (Pierre) et (Nicole) et (Jacques) et (Françoise)
- iv (Pierre et Nicole) et (Jacques et Françoise)

Dans UG p. 229, Montague caractérise une classe particulière de « modèles³⁵ » et il ajoute :

Let us add the assumption that *K* is a class of models for *L*. (The most important cases are those in which *K* is regarded as the class of logically possible models for *L*; among the conditions characterizing *K* might then appear the requirement that the « logical operations » and « logical words » of *L* receive their usual [logical] interpretations.).

À noter tout particulièrement que le *logical* entre crochets c'est nous qui l'avons inséré et c'est sans échappatoire dans le raisonnement de Montague. Or il saute aux yeux de celui qui veut étudier les langues naturelles en tant que langage objet susceptible du même type de connaissance que n'importe quel objet dans l'empirie, que les « logical words » dans les langues naturelles ne reçoivent pas leur « usual [logical] interpretations ». Ainsi et pour n'évoquer que le problème posé par (i) et qui devient patent pour la résolution de l'anaphore, on ne peut que faire le constat observationnel que (iv) peut être suivi de (v) alors que ce n'est pas le cas de (ii) et de (iii).

- v ces deux couples

Cela signifie qu'il n'est pas vrai que, contrairement à ce que donnerait LI en PTQ, (iii) et (iv) sont logiquement équivalents. Les « logical words », sans même considérer des exemples comme ceux qui suivent,

³⁴ Le terme *traduire* est ici utilisé dans son sens habituel et non dans son sens montaguien; on l'utilise donc avec le même sens que dans la citation de [Chambreuil 98a p.46] de la Section 2.

³⁵ Ces « modèles » ne sont pas les modèles de 5P, d'où notre utilisation des guillemets. Les « modèles » caractérisent l'univers de dénotation.

Éteins les lumières, ferme la porte et abandonne les lieux \neq Abandonne les lieux,
ferme la porte et éteins les lumières
Regarde ce film et tu meurs d'ennui

ne peuvent pas être traités dans les langues naturelles en leur donnant d'office l'interprétation que leurs « homologues » ont en logique. Ceci ne veut certainement pas dire que l'on ne peut pas avancer dans la connaissance du problème en le simplifiant, en se donnant un système d'observation qui ne considère que certains cas, en utilisant la capacité calculatoire de la logique pour décrire certains comportements, etc. Mais justement, pour faire tout ceci tout en évitant que le champ observé ne soit un terrain utilisé pour y exhiber l'utilisation de formalismes importés, il faut disposer d'un système d'observation systématique susceptible d'exprimer les observations dans un langage ad hoc.

L'exemple qui suit illustre cette même idée *a contrario*, c'est-à-dire qu'il montre l'espèce de désinvolture avec laquelle la question de l'observation est traitée. Soit la discussion des idées de [Pelletier 93, 94b] dans [Janssen 97 p. 444-445]. Considérons les exemples (i) et (ii) qui suivent, (respectivement numérotés (12) et (13) dans le texte cité) :

- i Kim believes that dentists usually need to hire an attorney
- ii Kim believes that tooth doctors commonly require the professional services of a lawyer

Ce qui nous intéresse ici, ce n'est pas de prendre position par rapport au fond du problème discuté (en gros s'il y a ou non des synonymes), mais d'observer comment [Janssen 97 p. 445] se débarrasse de la question de l'observation :

However, it [is] easy to make up some story in which Kim believes the embedded sentence in [i], but not the one in [ii].

Le point à remarquer dans cette discussion est qu'il se peut qu'il soit facile de forger une histoire anecdotique pour faire passer un exemple ponctuel, mais c'est incroyablement plus difficile de forger un langage d'observation, intersubjectivement et systématiquement utilisable par l'Observateur et permettant d'organiser les observations dans, p. ex., le domaine des lectures *de re* et *de dicto* ³⁶.

L'adoption d'un langage de calcul logique, p. ex. LI, est la deuxième idée avancée pour le développement de la sémantique dans 5P. Cette idée, en quelque sorte, complète et précise la précédente dans une direction déterminée, parmi les deux grandes directions possibles.

³⁶ Dans [Bès 85] a été esquissée cette possibilité : dès qu'on généralise le système d'observation habituel on retrouve des distinctions entre ces deux lectures dans des structures syntaxiques qui vont très au-delà de celles habituellement citées dans la présentation des cas d'école.

En examinant l'Idée 3 de Montague présentée dans la Section 2 et analysée dans la Section 3.1, celle qui fixe à la sémantique comme objectif la caractérisation des notions d'énoncé vrai et de conséquence, nous avons volontiers admis que cet objectif du projet de Montague était, en tant que choix mais non en tant que nécessité inéluctable, possible et intéressant, même si l'argumentation connue pour le justifier comme exclusif est plutôt légère. C'est-à-dire que l'on veut laisser en 5P la porte ouverte pour aboutir à des représentations sémantiques non calculables, des représentations dont l'objectif n'est pas de rendre compte de la fausseté et de la vérité des énoncés, des représentations qui relèvent, p. ex., de la sémantique développée par Mel'čuk et ses collaborateurs (cf. note 28).

Mais supposons que l'on reste à l'Idée 3 de Montague en tant que choix possible sur l'objectif à fixer à la sémantique. Il n'est pas évident que tout ce qui est aujourd'hui observable systématiquement soit aujourd'hui (ou toujours?) représentable dans un langage calculable. Il se peut qu'il y ait des limites. Les explorer serait un apport de première importance. Pour l'immédiat il semble raisonnable d'introduire comme limite à l'observation que celle-ci soit exprimable dans les formules d'une logique puissante. Et LI, forgée par Montague, semble une candidature naturelle et un outil puissant³⁷, bien qu'il y en ait d'autres (DRT, p. ex.) susceptibles de jouer un rôle analogue. Cela conduit à fixer la troisième idée pour le développement de la sémantique dans 5P.

En nous étant démarqué de la compositionnalité^{hu} en tant que commandement méthodologique à respecter, nous n'entendons pas nous cantonner dans la spécification des représentations sémantiques adjointes en bloc aux structures syntaxiques. Nous cherchons une association « compositionnelle » au sens de systématisme de la traduction par les Fonctions sémantiques. Celles-ci reçoivent en entrée des structures d'un certain type – le graphe constitué par l'ensemble de paires de fléchage – et donnent en sortie les formules d'une logique quelconque et, tout particulièrement, les formules de LI. En bref, nous restons à la *compositionnalité*^{strad} de la Section 3.2.

En résumé, pour développer la sémantique de 5P, il s'agirait :

- d'observer sémantiquement les énoncés dans la mesure de ce qui est exprimable par LI – ce qui est déjà énorme – et d'exprimer ces observations portant sur une sous-classe d'énoncés au moyen d'un langage d'observation ;
- de décrire ces énoncés moyennant les Propriétés et/ou les Projections, de manière à obtenir les modèles correspondants, avec, notamment l'ensemble des paires de fléchage et la représentation sémantique des entrées du Lexique instanciées dans les modèles ;
- de spécifier les Fonctions sémantiques, celles qui vont prendre en entrée les modèles avec leur fléchage et les sémantiques des entrées et qui vont donner les représentations sémantiques, c'est-à-dire, p. ex., des formules de LI. La production des représentations sémantiques peut se concevoir de deux manières différentes : soit directement, dans les formules d'une logique quelconque, soit en passant par un langage pré-sémantique intermédiaire.

³⁷ C'est ce que l'on fait dans [Morrill 94] en utilisant LI comme cible de la traduction de la syntaxe.

NOTES COMPLÉMENTAIRES

Note N° 1. Pour prouver que l'affirmation du texte de la Section 7.2 in fine, que nous reproduisons par la suite

Comme indication que les formules I à IX ne sont pas erronées, on peut faire le constat suivant : on ne connaît aucun développement, amélioration, modification ou adaptation de la Grammaire de Montague avec incidence sur l'algèbre de la syntaxe utilisée pour décrire les langues naturelles ayant réussi à produire une description de celles-ci incorporant les opérations syntaxiques formelles et computables – et non pas des manipulations du type de celles employées dans pSLF – d'une algèbre susceptible d'être associée par un homomorphisme- u à LI tel que celui défini dans UG et ceci dans un domaine observationnel incorporant le domaine visé dans pSLF.

concerne effectivement les travaux cités dans la note 25, il faudrait reprendre un par un ces travaux et signaler en quoi chacun est concerné par l'affirmation, ce qui serait très fastidieux. Mais on peut montrer sur deux exemples significatifs comment illustrer l'affirmation générale précédente.

Le couple-pSLF $\langle V_{16.n}, F_{10.n} \rangle$ est nécessaire pour exprimer, parmi d'autres, l'ambiguïté de portée des quantificateurs. Le problème posé par cette dernière question peut se résumer ainsi :

- Dans la syntaxe LI les formules $\forall x \exists y \alpha$ et $\exists y \forall x \alpha$ ne sont pas équivalentes.
- Dans la syntaxe des langues naturelles on a des structures
(SN₁) α (SN₂)

telles que si elles sont traduites sur des formules logiques avec des quantificateurs – que nous notons respectivement ici Q_1 et Q_2 –, l'observation indique la présence d'une ambiguïté, c'est-à-dire que la structure syntaxique précédente doit pouvoir se traduire par les deux formules suivantes

$$\begin{array}{l} Q_1 Q_2 \alpha \\ Q_2 Q_1 \alpha \end{array}$$

La clef du problème est que l'ordre dans la syntaxe des langages logiques est pertinent dans ce cas pour exprimer une différence d'interprétation, alors que cela n'est pas le cas en langue naturelle.

La syntaxe de PTQ a impérativement besoin de $\langle V_{16.n}, F_{10.n} \rangle$ pour exprimer l'ambiguïté de LN. Ce couple pSLF fonctionne moyennant l'artifice d'obtenir deux dérivations différentes en plaçant SN₁ et SN₂ dans deux ordres différents. Lorsqu'on a recours à l'ordre SN₂ SN₁, une manipulation viendra pour faire disparaître le he_0 , qui, lui, a été

placé à la bonne place, et qui sera remplacé par SN₂. Ainsi pour pouvoir obtenir la portée large de l'existentiel pour SN₂, on aura en syntaxe les deux lignes :

SN₂ SN₁ ... he_j ...
SN₁ ... SN₂

Tout cela hors formalisation véritable et hors computabilité. On a donc essayé de modifier cette manière de faire, p. ex. dans [Cooper 83]. Le *Cooper storage* est en définitive une mémoire qui conserve « au frigo » les SN pour les sortir au moment adéquat. Comme il est dit dans [Cooper 83 p. 55]

[...] the storage technique involves putting an NP interpretation on ice for a while until you have interpreted enough of the tree to represent the scope you want to give the NP.

Or, ce n'est pas cette technique en soi qui est discutée ici. La question est que l'utilisation de la mémoire nécessite des opérations – notamment de mise en mémoire et de récupération – que l'on ne voit pas du tout associées par un homomorphisme-*u* à LI. De ceci [Cooper 83 p. 12] est parfaitement conscient :

[...] if one defines compositionality as what one is allowed to do in interpreting, say, Montague's intensional logic, then the enrichment of interpretation which I will undertake in chapters III and IV (involving storage and substitution) will not provide a compositional semantics. However, it is still the case that the semantics is defined recursively on the syntax, so in this sense it is compositional.

Cela veut dire que l'on respecte la compositionnalité en se donnant de compositionnalité une définition autre que celle de l'homomorphisme-*u* (cf. la définition donnée dans la Note 2 ci-dessous).

Autre exemple qui va exactement dans le même sens : [Partee 79a] (cf. sa présentation dans [Chambreuil 89 p. 178-195], où, cependant, le problème qui nous occupe ici n'est pas abordé). Dans les propositions de [Partee 79a], il y a une typification des opérations primitives en opérations :

- i de concaténation et adjonction
- ii de substitution (« substitution simple » et « substitution partout »)
- iii de spécification de propriétés [les propriétés ressemblent comme deux gouttes d'eau aux traits des grammaires à traits]
- iv de duplication de propriétés.

Par ailleurs, on suppose une composante morphologique proposée dans [Partee 79b]. Par rapport à la question qui nous occupe, il faut remarquer :

- qu'on ne voit pas la possibilité de définir un homomorphisme-*u* entre opérations de la syntaxe et opérations de LI ou d'une algèbre dérivée de LI ;
- que le recours à une composante morphologique est une conséquence qui semble incontournable pour traiter la LN, et qui vient au fond de la formule VI de la Section 7.2; et sur ce point aussi, on ne voit pas l'homomorphisme-*u* avec LI ;

- qu'on ne voit pas la computabilité possible pour au moins l'opération de « substitution partout » (cf. [Partee 79a p. 70-71]) ;
- que [Partee 79a p. 53] définit une compositionnalité qui n'est pas celle de UG. Elle est définie sur les règles de la syntaxe qui, pour ainsi dire, agglutinent des opérations (p. ex. de concaténation, substitution ou autre avec des opérations de spécification des propriétés) et non sur des opérations simples. Par ailleurs, les opérations incorporées aux règles « may be as simple as concaténation or as complex as a transformational operation ». Cela veut dire qu'on y retrouve la question de la non-computabilité.

Voilà le type d'analyse qu'implacablement on doit pouvoir faire sur toute tentative passée ou à venir essayant d'« améliorer » la syntaxe pseudo-formelle de PTQ. Si on la rend explicitement formelle et computable, alors les exigences sur l'homomorphisme- μ pour aboutir à l'interprétation induite via LI ne sont pas respectées : c'est que la naturalité de la langue naturelle est à l'œuvre.

Note N° 2. Comme il a été signalé dans le texte, la non-empiricité de la compositionnalité de la Grammaire de Montague avait déjà été remarquée. Elle peut être étendue : Montague n'a jamais pris la peine de bien préciser les observations qu'il était censé décrire. Dans les premières lignes de PTQ (p. 247) il se limite à observer que :

[...] the fragment has been made as simple and restricted as it can be while accommodating all the more puzzling cases of quantification and reference with which I am acquainted.

Plus généralement, il n'est pas injuste de dire que Montague et ses supporters n'ont pas eu comme souci principal le fait de pratiquer la linguistique comme science empirique, ce qui risque de conduire à une impasse, ou stérilité, comme dit [van Benthem 86 p. 180] :

A modern semantic research program like Montague Grammar runs precisely the same risk of sterility, if its practitioners do not start adding falsifiable claims soon.

Malgré cela, il n'est pas faux de dire qu'il existe un consensus relativement explicite sur le fait que les représentations sémantiques que l'on veut associer aux énoncés de LN doivent rendre compte de quelque chose au niveau de l'observation, et ce quelque chose est ce qu'on nomme des « intuitions » sur les énoncés. Comme signalé par [Cooper 83 p. 9] (et cf. aussi [Dowty & al. 81 p. 1])

For the better or worse, with semantics as with syntax, the linguist has as primary data the intuitions of native speakers.

Montague, avec son Idée 2, écarte la possibilité d'accéder à des intuitions sur la représentation de la syntaxe; il en est de même dans 5P. On suppose donc la possibilité d'explicitement des Observations sur la sémantique, même si ce n'est pas simple. Comme dit [Cooper 83 p. 8] cela « is often a subtle and complex practical problem ». Tellement subtil et complexe qu'on chercherait en vain dans tous les écrits sur la sémantique formelle un travail qui ne serait pas une *méthode*, dans le sens de recettes à suivre, pour obtenir comme

résultat des observations, mais qui proposerait des conventions permettant de systématiser des observations sémantiques et leur expression dans un langage de représentation ad hoc. On s'éviterait ainsi d'introduire des anecdotes ad hoc pour justifier le bien fondé de telle ou telle représentation sémantique.

Il est impossible dans le cadre de ce document d'avancer des propositions aidant à systématiser les observations sur la sémantique. Nous nous limitons à montrer deux exemples de sa justification, l'un qui ne conduit pas vers une sémantique testable dans l'empirie et l'autre qui le fait.

L'exemple qui ne conduit pas à la systématisation souhaitée illustre toutes les situations où, de manière déguisée, on introduit dans l'observation des contraintes qui viennent du système d'hypothèses censées rendre compte de ces mêmes observations.

Soit, pour illustrer cette manière de faire, la compositionnalité telle qu'elle est envisagée par [Cooper 53 p. 5]. Elle est définie ainsi :

When a rule in the syntax tell us how to put phrases together to make larger phrases, the semantics has a corresponding rule which tells us how to combine the meanings of the phrases we are putting together to make the meaning of the larger phrase. This technique has come to be known as the principle of compositionality [...]

Par ailleurs [Cooper 83 p. 6] observe que les expressions de l'anglais *a man*, *every woman*, *no child* sont des phrases bien formées de l'anglais et que, si on veut leur donner une traduction en logique de premier ordre, on trouve comme difficulté le fait que

[...] the predicate calculus sentence corresponding to *every woman runs* is very different to the syntact structure of the English sentence.

Et il voit deux solutions :

- soit refaire la structure de la syntaxe de manière à l'accommoder à la syntaxe de la formule en logique des prédicats ;
- soit « to find some appropriate model-theoretic objects to interpret the noun-phrases that are given to us by the syntax of English » [Cooper 83 p. 6-7].

Naturellement il choisit la deuxième solution en adoptant les propositions de Montague.

Il est à noter que ce que l'on gagne ici est la satisfaction de la compositionnalité de type règle à règle. Aucune avancée n'est enregistrée sur la description des observations portant sur chaque syntagme nominal ou sur les structures qui vont se créer par l'incorporation de ces syntagmes nominaux au reste de la phrase. On ne valide ni ne teste aucune description en sémantique formelle par rapport à l'observation sur les énoncés, mais par le fait qu'elle satisfait une exigence que l'on avait bien décidé d'imposer à cette sémantique formelle. Ainsi si on suppose une seule règle en syntaxe associant un déterminant à un nom, ou

intégrant un syntagme nominal à une phrase, on veut qu'il y ait une seule règle correspondante en logique, et comme on ne trouve pas cette association règle à règle, on décide de changer de logique pour pouvoir le faire. Cf. le même raisonnement dans [Dowty & al. 81 p. 106-107]. On ne voit pas l'empiricité dans ce type de raisonnement.

Comme illustration positive de ce que peut être une sémantique testable empiriquement, on reprend la question de l'interprétation à donner aux déterminants dans le syntagme nominal. Supposons la structure

[Det N (M)] SN SV

avec :

- M un modifieur optatif
- SV un syntagme verbal quelconque

Acceptons que SN a une dénotation, que si M n'est pas utilisé, cette dénotation sera notée SN' , alors que si M est utilisé, elle sera notée SN'' . Acceptons aussi que $SN'' \subseteq SN'$. Comme exemple, on peut observer que les objets dénotés par (i) – avec SN'' – sont un sous-ensemble des objets dénotés par (ii) – avec SN' .

- i les brioches de fabrication artisanale
- ii les brioches

Il se trouve que selon le Det utilisé, l'utilisation de SN' à la place de SN'' peut conserver ou non les valeurs de vérité dans la structure syntaxique présenté plus haut. Par exemple, soit le déterminant *quelques* et les deux phrases suivantes; si (i) – avec SN'' – est un énoncé vrai, (ii) – avec SN' – le sera aussi, mais le contraire n'est pas vrai.

- i Quelques brioches de fabrication artisanale étaient bonnes
- ii Quelques brioches étaient bonnes

Le déterminant *tous/tes* présente un comportement inversé; si (ii) – avec SN' – est un énoncé vrai, (i) – avec SN'' – le sera aussi, mais le contraire n'est pas vrai.

- i Toutes les brioches de fabrication artisanale étaient bonnes
- ii Toutes les brioches étaient bonnes

On va appeler *monotoniques croissants* les déterminants comme *quelques* et *monotoniques décroissants* les déterminants comme *tous/tes*.

À partir d'observations de ce type on peut classer les déterminants, et ce qui est encore plus intéressant, établir des principes généraux portant sur le type de structures syntaxiques associables aux différents déterminants ainsi classés (cf. [Partee & al. 90 chap. 14] d'où on a tiré les exemples présentés ici, et [van Benthem 86, 1ère Partie]). P. ex., un de ces

principes exprime des restrictions sur l'utilisation de la négation portant sur les syntagmes nominaux. D'un de ces principes découlerait :

an explanation for the fact that not all NPs can be negated, e.g., acceptable external negation [négations portant sur le SN] are :

(i) not every man

...

whereas unacceptable are :

(4) *not few man".

On a ainsi un authentique principe empirique sur les LN. Tellement authentique qu'il est possible de le tester et de le mettre en difficulté avec des contre-exemples. Le principe général voudrait que la négation ne puisse pas porter sur les déterminants monotoniques décroissants, parmi lesquels on range *few*. Or le principe est vérifié dans (i) en français qui suit, mais il ne l'est pas dans l'espagnol (ii)

i *Pas peu ont rejeté la proposition

ii No pocos rechazaron la propuesta

Toute hypothèse empirique sur les LN est infiniment plus risquée que la présentation des formalismes validés par leur utilisation dans quelques exemples choisis, et l'exemple précédent n'est pas destiné à évaluer les hypothèses empiriques sur les déterminants mais plutôt à signaler la possibilité d'un développement effectivement empirique de la sémantique des langues naturelles.

Références

- EFL
Montague, Richard "English as a Formal Language". Dans [Thomason 74 p. 188-221].
- PTQ
Montague, Richard "The Proper Treatment of Quantification in Ordinary English". Dans [Thomason 74 p. 247-270].
- UG
Montague, Richard "Universal Grammar". Dans [Thomason 74 p. 222-246].
- [Amsili & Bras 98]
Amsili, Pascal & Myriam Bras "DRT et compositionnalité". Dans [Nazarenko 98a p.131-160].
- [Auroux 93]
Auroux, Sylvain *La logique des idées*. Montréal..., Bellarmin..., 1993.
- [Bach 79]
Bach, Emmon "Montague Grammar and Classical Transformational Grammar". Dans [Davis & Mithun 79 p. 3-49].
- [Bès 85]
Bès, Gabriel G. "Référence, ambiguïté et paraphrase". Dans Catherine Fuchs [éd.] *Aspects de l'ambiguïté et de la paraphrase dans les langues naturelles*. Berne..., Peter Lang, 1985, p. 167-206.
- [Bès 99]
Bès, Gabriel G. "La phase verbale noyau en français". Dans G.A.R.S., N°15 (1999), p.273-358.
- [Bès & Hagège 01]
Bès, Gabriel G. & Caroline Hagège "The 5P Paradigm" [en cours d'achèvement]
- [Bunge 69]
Bunge, Mario. *La investigación científica; Su estrategia y su filosofía*. Barcelona, Ariel, 1969.
- [Chambreuil 89]
Chambreuil, Michel *Grammaire de Montague. Langage, traduction, interprétation*. Clermont-Fd, Adosa, 1989.
- [Chambreuil 91]
Chambreuil, Michel "Expressions nominales, sémantique intensionnelle, sémantique situationnelle". Dans *Cahiers du DLSL*, (10, 1991), p. 9-48.
- [Chambreuil 98]
Chambreuil, Michel [dir] *Sémantiques*. Paris, Hermes, 1998.
- [Chambreuil & al 98]
Chambreuil, Michel, Abdeljabbar Ben Gharbia & Pablo Gamallo Otero "Variations sur la compositionnalité Montaguienne". Dans [Nazarenko 98a p. 35-65].
- [Chambreuil & Pariente 90]
Chambreuil, Michel & Jean-Claude Pariente *Langue naturelle et logique; La sémantique intensionnelle de Richard Montague*. Berne, Peter Lang, 1990.

- [Cooper 83]
Cooper, Robin *Quantification and syntactic theory*. Dordrecht... , D. Reidel Publishing Company, 1983.
- [Davis & Mithun 79]
Davis, Steven & Marianne Mithun [eds.] *Linguistics, Philosophy, and Montague Grammar*. Austin..., University of Texas Press, 1979.
- [Dowty 79]
Dowty, David R. *Word Meaning and Montague Grammar; The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*. Dordrecht..., D. Reidel Publishing Company, 1979.
- [Dowty 82]
Dowty, David "Grammatical Relations and Montague Grammar". Dans [Jacobson & Pullum 82 p. 79-130].
- [Dowty & al 81]
Dowty, David R., Robert E. Wall & Stanley Peters *Introduction to Montague Semantics*. Dordrecht..., D. Reidel Publishing Company, 1981.
- [Feyerabend 75]
Feyerabend, Paul *Against Method*. London, New Left Books, 1975.
- [Feyerabend 96]
Feyerabend, Paul *Tuer le temps*. Paris, Éditions du Seuil, 1996.
- [Friedman & Warren 78]
Friedman, J. & D.S. Warren "A parsing method for Montague Grammars". Dans *Linguistics and Philosophy*, 2(1978).
- [Gazdar & al 85]
Gazdar, Gerald, Ewan Klein, Geoffrey Pullum & Ivan Sag *Generalized Phrase Structure Grammar*. Oxford, Basil Blackwell, 1985.
- [Granger 92]
Granger, Gilles-Gaston *La vérification*. Paris, Éditions Odile Jacob, 1992.
- [Groenendijk & al 83]
Groenendijk, J.A.G., T.M.V. Janssen & M.B.J. Stokhof [eds.] *Formal Methods in the Study of Language*. Amsterdam, Mathematisch Centrum, 1981.
- [Halvorsen & Ladusaw 79]
Halvorsen, Per-Kristian & William A. Ladusaw "Montague's 'Universal Grammar': An Introduction for the Linguist". Dans *Linguistics and Philosophy*, 3 (1979), p. 185-223.
- [Jacobson & Pullum 82]
Jakobson, Pauline & Geoffrey K. Pullum [eds.] *The Nature of Syntactic Representation*. Dordrecht..., D. Reidel Publishing Company, 1982.
- [Janssen 97]
Janssen, T.M.V. "Compositionality". Dans [van Benthem & ter Meulen 97 p. 417-473].
- [Kamp & Reyle 93]
Kamp, Hans & Uwe Reyle *From Discourse to Logic*. Dordrecht..., Kluwer Academic Publishers, 1993.
- [Kuhn 70]
Kuhn, Thomas *The structure of Scientific Revolutions*, 2nd edition. Chicago, University of Chicago Press, 1970.

- [Landsbergen 83]
Landsbergen, Jan "Adaptation of Montague Grammar to the requirements of parsing". Dans [Gronendijk & al. 83 p. 399-420].
- [Lewis 70]
Lewis, David "General Semantics". dans *Synthese*, 22 (1970), p. 18-67. [Cité dans [Partee 97]].
- [Mel'čuk 88]
Mel'čuk, Igor A. *Dependency Syntax : Theory and Practice*. State University of New York Press, 1988.
- [Mel'čuk & Pertsov 87]
Mel'čuk, Igor A. & Nikolaj V. Pertsov *Surface syntax of English; A formal model within the meaning-text framework*. Amsterdam..., John Benjamins Publishing Company, 1987.
- [Morrill 94]
Morrill, Glin V. *Type Logical Grammar; Categorical Logic of Signs*. Dordrecht..., Kluwer Academic Publishers, 1974.
- [Morris 38]
Morris, C. N. *Foundations of the theory of signs*. Chicago, 1938.
- [Nazarenko 98a]
Nazarenko, Adeline *Compositionnalité*. t.a.l., 39 (98).
- [Nazarenko 98b]
Nazarenko, Adeline "Présentation". Dans [Nazarenko 98a p. 3-7].
- [Pariante 85]
Pariante, Jean-Claude *L'analyse du langage à Port-Royal. Six études logico-grammaticales*. Paris, Les Éditions de Minuit, 1985.
- [Partee 79a]
Partee, Barbara H. "Constraining Transformational Montague Grammar : A Framework and a Fragment". Dans [Davis & Mithun 79 p. 51-101].
- [Partee 79b]
Partee, Barbara H. "Montague Grammar and the well-formedness constraint". Dans F. Heny & H.Schenelle [eds.] *Syntax and Semantics*, vol 10. New York, Academic Press, 1979.
- [Partee 84]
Partee, Barbara H. "Compositionality". Dans F. Landman & F. Veltman *Varieties of Formal Semantics*. Dordrecht, Foris, 1984, p. 282-311.
- [Partee 95]
Partee, Barbara H. "Lexical Semantics and Compositionality". Dans L. Gleichman & M. Liberman *An invitation to Cognitive Science: Language*. Cambridge (Mass), MIT Press, 1995, p. 311-360.
- [Partee 97]
Partee, Barbara H. with H.L.W. Hendriks "Montague Grammar". Dans [van Benthem & ter Meulen 97 p. 5-91].
- [Partee & al 90]
Partee, Barbara H., Alice ter Meulen, Robert E. Wall *Mathematical Methods in Linguistics*. Dordrecht..., Kluwer Academic Publishers, 1990.
- [Pelletier 93]
Pelletier, F.J. "Some issues involving internal and external semantics". Dans J. Macnamara & C. Reges [eds.], Oxford, Oxford University Press, 1993 [Cité dans Janssen 97].

- [Pelletier 94]
Pelletier, F.J. "The principle of semantic compositionality". dans *Topoi*, 13 (1994), p. 11-24 [Cité dans Janssen 97].
- [Peters & Ritchie 73]
Peters, P. S. & R.W.Ritchie "On the generative power of transformational grammars". Dans *Information Sciences*, 6 (1973), p.49-83.
- [Poirier & al. 98]
Poirier, Christèle, Yann Mathet & Patrice Enjalbert "La compositionnalité à l'épreuve des faits, à travers un projet de compréhension automatique de constats d'accidents". Dans [Nazarenko 98a p.99 -129].
- [Sokal & Bricmont 97]
Sokal, Alan & Jean Bricmont *Impostures intellectuelles*. Paris, Éditions Odile Jacob, 1997.
- [Thomason 74]
Thomason, Richmond H. [ed.] *Selected papers of Richard Montague*. New Haven..., Yale University Press, 1974.
- [Trouilleux & al 00a]
Trouilleux, François, Gabriel G. Bès & Éric Gaussier "An Evaluation of Inter-Annotator Agreement in the Observation of Anaphoric and Referential Relations". Dans *Proceedings of the Third International Conference on Discourse Anaphora and Anaphor Resolution (DAARC 2000)*. Lancaster University, UK, 2000.
- [Trouilleux & al 00b]
Trouilleux, François, Éric Gaussier, Gabriel G. Bès & Annie Zaenen "Coreference Resolution Evaluation Based on Descriptive Specificity". Dans *Proceedings of the Second International Conference on Language Resources and Evaluation (LREC 2000)*. ELRA, Athènes (Grèce), 2000.
- [van Benthem & ter Meulen 97]
J. van Benthem & A. Ter Meulen [eds] *Handbook of Logic and Language*. Amsterdam..., Elsevier... , 1997.
- [van Benthem 86]
van Benthem, Johan *Essays in Logical Semantics*. Dordrecht..., D. Reidel Publishing Company, 1986.
- [Warren & Friedman 82]
Warren, D.S. & J. Friedman "Using semantics in non-context-free parsing of Montague Grammar". Dans *American Journal of Computational Linguistics*, 8 (1982).
- [Zadrozny 94]
Zadrozny, Z. "From compositional to systematic semantics". Dans *Linguistics and Philosophy*, 12 (1994), p. 95-131 [Cité dans [Janssen 97]].

Annexe

Grammaire de Montague et ambiguïté

Sommaire

1	Questions sur le traitement de l'ambiguïté dans la Grammaire de Montague.....	3
2	Ambiguïté d'assignation catégorielle dans UG	4
	2.1 Possibilité formelle	
	2.2 Répercussions systémiques	
	2.2.1 L'ambiguïté syntaxique	
	2.2.2 L'ambiguïté sémantique	
	2.2.3 L'ambiguïté de traduction	
3	Ambiguïté d'assignation catégorielle dans PTQ.....	17
	3.1 Possibilité formelle	
	3.2 Répercussions systémiques	
4	Réponses aux questions sur le traitement de l'ambiguïté dans la Grammaire de Montague.....	25

1 Questions sur le traitement de l'ambiguïté dans la Grammaire de Montague*

On sait que la notion d'ambiguïté est facile à comprendre intuitivement : on a un même objet, objet conçu en tant qu'objet perceptible, phonique ou graphique, qui admet plusieurs « interprétations ». Le problème est d'expliciter ce que sont ces « interprétations » et/ou dans quelles conditions se réalise l'admission de ces « interprétations ». Et pour répondre à ces questions, on ouvre la boîte de Pandore qui enferme les questions sur le référent, l'encyclopédie, le rapport syntaxe-sémantique, etc.¹ : encore une notion « auberge espagnole d'autrefois » dans le domaine de l'étude des langues naturelles.

Aussi bien UG que PTQ abordent la question de l'ambiguïté dans des cadres plus précis que ceux de la question générale de l'ambiguïté. Mais ces cadres ne sont pas les mêmes dans les deux travaux, bien qu'ils aient des points communs entre les deux. Or pour aborder cette question nous retrouvons de plein fouet le problème du flottement terminologique déjà signalé dans la Section 1.1 du document « Empiricité en linguistique et Grammaire de Montague... » (document ci-avant, désormais EGM), flottement concernant maintenant le terme *catégorie*. Nous utilisons ici une terminologie désambiguïsée, mais avec une notation qui s'éloigne peu de la notation montaguienne ; par ailleurs nous insérons dans le texte ou en note des commentaires qui devraient permettre de suivre la discussion par rapport aux textes originaux.

Deux questions similaires peuvent être posées dans le cadre aussi bien d'UG que de PTQ. Elles portent sur les deux points suivants.

- Est-ce qu'en syntaxe on peut avoir des expressions graphiques, élémentaires ou non, associées à plus d'une catégorie?
- Si la réponse à la question précédente est oui, quelles sont les conséquences que cela entraîne pour les autres pièces du système de la Grammaire de Montague, en particulier pour la traduction en LI et/ou l'interprétation de la syntaxe dans un univers de dénotation?

* Pour les références bibliographiques de cette Annexe, cf. la Section "Références" du document EGM ci-avant.

¹ Pour une présentation d'ensemble des problèmes soulevés par la notion générale d'ambiguïté, cf. [Le Goffic 81]. On sait que l'on peut être incommodé par le fait de parler des "ambiguïtés" des textes quand elles ne gênent en rien le lecteur puisque celui-ci ne s'aperçoit même pas qu'elles existent. On n'essaye pas ici de traiter ces questions : on reste dans le cadre étroit de l'ambiguïté telle qu'elle est vue dans le système de Montague (UG et PTQ).

La première question sera dite ici « question de la possibilité formelle de l'ambiguïté d'assignation catégorielle »; la deuxième « question des répercussions systémiques de l'ambiguïté d'assignation catégorielle ».

Du fait même qu'en PTQ il n'y a pas, contrairement à UG, de « langage non ambigu », ces deux questions peuvent et doivent se poser différemment dans les deux cadres. C'est ce que nous ferons dans les Sections 2 et 3 qui suivent pour respectivement UG et PTQ.

2 Ambiguïté d'assignation catégorielle dans UG

Pour poser la première question, dans le cadre de UG, on suppose ce qui suit :

- i La définition dans UG de « langage non ambigu »
- ii Les contraintes afférentes assurant la non ambiguïté
- iii La classe \underline{C} d'ensembles indexés d'expressions générées par un « langage non ambigu » telle que
 $\underline{C} = \{C_1, \dots, C_n\}$ avec C_i un ensemble d'expressions graphiques, élémentaires ou non, et $i \in \Delta$, i étant un *indice catégoriel* qui sera noté par une lettre grecque en minuscules.
- iv La propriété d'unicité d'assignation catégorielle est définie ainsi :
 Étant donné :
 - le « langage non ambigu » L avec ses entités associées \underline{C} et Δ L et \underline{C} satisfont chacun la propriété d'unicité d'assignation catégorielle si étant donné :
 - $C_\delta, C_{\delta'} \in \underline{C}$
 - $\delta, \delta' \in \Delta, \delta \neq \delta'$
 on a alors $C_\delta \cap C_{\delta'} = \{\}$

Dans le cadre posé par (i) à (iv) on peut expliciter la première question ainsi :

Est-ce que (i), (ii), (iii) vont assurer que tout « langage non ambigu » va satisfaire la propriété d'unicité d'assignation catégorielle?

Dans la suite on va noter « langage non ambigu » par $L \sim A$, et on utilisera $L \sim AU$ pour noter un « langage non ambigu » qui satisfait la propriété d'unicité d'assignation catégorielle, et $L \sim A \sim U$ pour noter un « langage non ambigu » qui ne la satisfait pas. La question précédente peut se reformuler donc ainsi : est-ce que les définitions de la Grammaire de Montague autorisent formellement l'existence d'un objet tel que $L \sim A \sim U$?

La deuxième question, la question des répercussions systémiques de l'ambiguïté d'assignation catégorielle, toujours dans le cadre de UG, peut s'expliciter de la manière suivante. En plus de supposer qu'un objet tel que $L \sim A \sim U$ a une existence formelle possible, supposons :

- l'utilisation de parenthèses indexées introduites par chaque opération F_γ (cf. ci-dessous Section 2.1 le quintuplet définissant « langage ») ;
- la relation R d'effacement des parenthèses indexées ;
- l'utilisation de R pour obtenir un « langage » éventuellement ambigu à partir de $L \sim A$.

Dans le cadre ainsi posé, la deuxième question devient :

Étant donné $L_1 \sim AU$ et $L_2 \sim A \sim U$, quelles sont les différences que l'on doit trouver par rapport

- à l'ambiguïté syntaxique, étant donné L'_1 et L'_2 tels que

$$L'_1 = \langle \langle L_1 \sim AU \rangle, R \rangle$$

$$L'_2 = \langle \langle L_2 \sim A \sim U \rangle, R \rangle,$$
- à l'interprétation de L'_1 et de $\langle L_1 \sim AU \rangle$, et de L'_2 et de $\langle L_2 \sim A \sim U \rangle$,
- à la traduction de L'_1 et de $\langle L_1 \sim AU \rangle$, et de L'_2 et de $\langle L_2 \sim A \sim U \rangle$ dans LI ?

La deuxième question ainsi posée exige un *caveat*. UG n'exige pas formellement l'utilisation de parenthèses, indexées ou pas². Les parenthèses sont un artifice d'écriture utilisé dans beaucoup de langages dits logiques, bien qu'elles y soient utilisées dans une version « soft », c'est-à-dire sans indices. Elles sont commodes, car elles ne sont pas utilisées pour enfermer les expressions graphiques élémentaires, ce qui permet de satisfaire facilement une des deux contraintes sur la non-ambiguïté des $L \sim A$ (cf. ci-dessous Section 2.1). Elles permettent de coder, dans l'objet généré, l'opération particulière utilisée pour générer cet objet. Par ailleurs l'opération d'effacement est simple, et, selon la direction qu'on lui impose, elle peut accéder au statut de fonction : une suite parenthésisée n'est associée qu'à une seule suite dénudée de ses parenthèses par leur effacement.

En faisant les suppositions précédentes nous ne nous écartons donc pas de la tradition dans la matière et nous n'introduisons pas de bruit dans la discussion. Les limites de l'interrogation sur les répercussions systémiques de l'ambiguïté d'assignation catégorielle seront explicitées par la suite.

Les questions relatives à l'ambiguïté au sens large sont traitées dans UG autour de trois notions, formellement définies :

- Ambiguïté syntaxique, notion présente dans UG p. 226 dans la Section 2 « Syntax ».
- Ambiguïté sémantique, notion présente dans UG p. 227 dans la Section 3 « Semantics : Theory of Meaning ».
- Équivalence, notion présente dans UG p. 230 dans la Section 3 « Semantics : Theory of Reference ».

Dans UG n'est pas traitée la question de l'ambiguïté par rapport à la traduction sur LI. Mais comme le lien formel pour ce qui nous intéresse ici est le même entre syntaxe et

² Cf. [Chambreuil 89], en particulier l'exemple p. 54, pour une utilisation de parenthèses indexées. Ce n'est évidemment pas la seule manière des respecter la non-ambiguïté. [Dowty 2 & al. 81 p. 256] en utilise une autre qui a le grave inconvénient de nécessiter la notion d'arbre, absente de UG.

interprétation qu'entre syntaxe et LI, on peut extrapoler sans crainte la définition d'ambiguïté sémantique pour obtenir la notion formelle d'ambiguïté de traduction.

Il n'est pas nécessaire d'entrer ici dans (iii) – notion d'équivalence – qui nous ferait compliquer beaucoup l'organisation de l'univers de dénotation, sans pourtant ajouter quelque chose d'important à notre discussion. Par ailleurs nous allons supposer que la traduction de la « syntaxe » (au sens de UG donc) se fait sur LI et non sur un « langage » arbitraire³.

Les questions portant sur l'ambiguïté syntaxique, l'ambiguïté sémantique et l'ambiguïté de traduction, relèvent de la deuxième question, sur les répercussions systémiques; elles seront abordées dans les Sous-sections de la Section 2.2. Au préalable, dans la Section suivante, on aborde la première question, sur la possibilité formelle d'un $L \sim A \sim U$.

2.1 Possibilité formelle

Nous rappelons dans (i) qui suit le quintuplet de base qui va être utilisé dans la définition d'un « langage » et dans (ii) et (iii) les deux restrictions qui vont faire en sorte qu'un « langage » soit un « langage non ambigu »⁴.

- i $\langle A, F_\gamma, X_\delta, S, \delta_0 \rangle_{\gamma \in \Gamma, \delta \in \Delta}$
- ii « X_δ and the range of F_γ are disjoint whenever $\delta \in \Delta$ and $\gamma \in \Gamma$ » (UG p. 226).
- iii « for all γ and $\gamma' \in \Gamma$, all sequences a in the domain of F_γ and all sequences a' in the domain of $F_{\gamma'}$, if $F_\gamma(a) = F_{\gamma'}(a')$, then $\gamma = \gamma'$ and $a = a'$ » (UG p. 226).

(ii) établit ainsi que l'on ne peut pas obtenir par une opération F_γ ce que l'on a déjà introduit dans un ensemble indexé X_δ et (iii) que l'on ne peut pas obtenir un même objet dans A par deux opérations F_γ différentes.

Par la suite nous rapellons par ailleurs quelques points cruciaux pour comprendre le traitement de l'ambiguïté syntaxique dans UG.

1 Le A du quintuplet de (i) ci-dessus est un ensemble; par conséquent il n'y a pas deux éléments identiques dans A . Par ailleurs les objets appartenant à A sont des expressions graphiques, élémentaires ou non. Si elles sont des expression graphiques élémentaires, elles ne sont pas enfermées par des parenthèses. Si elles sont des expressions graphiques non élémentaires elles comportent, dans l'expression même, des parenthèses indexées par un indice des opérations F_γ , c'est-à-dire un indice appartenant à Γ (et non à Δ , cf. le point

³ $R = \text{Id}[\text{entité}]$ dans le cas de LI et nous n'avons pas un "langage ambigu" associé à LI différent de LI.

⁴ La traduction française "langage non ambigu" ne rend pas l'idée intuitive de l'original anglais "disambiguated language" qui dit bien ce qu'il dit : il s'agit d'un "langage" qui a été rendu non ambigu; mais nous conservons l'usage déjà établi. Nous ne donnons dans le texte que la partie des définitions et des restrictions qui concernent la discussion, en supposant connues toutes les autres (p.ex. : relation de F_γ avec A , la définition de l'indice distingué δ_0 , etc.). L'ensemble Δ dans le fragment de l'anglais décrit dans UG est un ensemble numérique mais aucune exigence sur Δ n'est introduite dans les définitions générales.

suivant). Enfin, on trouve dans A toutes les expressions graphiques élémentaires – c'est-à-dire les expressions de tous les X_δ – et toutes les expressions graphiques non élémentaires spécifiées par toutes les opérations F_γ , que ces suites soient bien formées (suites significatives dans la terminologie des travaux de Chambreuil) ou non. Ce sont les objets dans le S du quintuplet qui vont permettre de discriminer parmi les suites non élémentaires celles qui sont bien formées (cf. ci-dessous le point 3).

2 Une opération F_γ (qui est une fonction) encadre dans sa sortie ces arguments d'entrée par des parenthèses indexées par l'indice de l'opération, ces parenthèses ainsi indexées étant dans la sortie même de la fonction. Ainsi la sortie de (i) suivant – que l'on suppose être une opération de concaténation – est (ii) et non (iii)

- i $F_\gamma (\alpha, \beta) =$
- ii $(\alpha\beta)_\gamma$
- iii $\alpha\beta$

Le résultat d'une opération F_γ peut ainsi être réduit à une liste d'éléments concaténés, chaque élément de la liste étant soit une expression graphique élémentaire soit une parenthèse. Pour des raisons de lisibilité, nous allons écrire ces listes avec un blanc comme intervalle entre chaque élément; (ii) sera ainsi noté $(\alpha \beta)_\gamma$.

Un X_δ est un ensemble qui, lui, est indexé par un indice catégoriel (un indice de Δ), mais dont les éléments, des expressions graphiques élémentaires, ne sont ni indexés par δ ni parenthésés. On a ainsi, p. ex. (i) et non, respectivement, (ii) suivants :

- i $X_{nom} = \{\text{garçon, livre, ...}\}; \quad X_{art} = \{\text{le, ...}\}$
- ii $X_{nom} = \{\text{garçon}_{nom}, \text{livre}_{nom}, ... \}; \quad X_{art} = \{\text{le}_{art}, ... \}$

De telle manière que, comme le α et le β de l'exemple précédent sont respectivement instanciés par des éléments de X_{nom} et de X_{art} , on aurait, à partir de F_γ , de X_{nom} et de X_{art} :

$$A = \{\text{garçon, livre, le, (garçon livre)}_\gamma, (\text{ le le })_\gamma, (\text{ le livre })_\gamma, (\text{ le garçon })_\gamma, \dots\}$$

À noter donc que tous les éléments de A ne sont pas bien formés.

3 Ce sont les objets dans S , malheureusement nommés parfois *règles* et dits CBF (couples de bonne formation) dans EGM qui vont filtrer les éléments de A que l'on ne veut pas considérer comme étant bien formés⁵; dans l'exemple précédent on devrait donc filtrer $(\text{ garçon livre })_\gamma$ et $(\text{ le le })_\gamma$. On continue ici l'usage terminologique de CBF et on présente les CBF avec une notation qui se veut plus intuitive et simple que celle de UG. Un CBF incorpore une opération F_γ , une liste d'arguments qui sont des indices de Δ , et donne en résultat un indice de Δ :

$$\langle F_\gamma, \langle \text{indice}_1, \dots, \text{indice}_n \rangle, \langle \text{indice résultat} \rangle \rangle$$

⁵ UG p. 225 ne donne pas de nom officiel aux éléments de S , en les caractérisant simplement comme des « sequences » dont la syntaxe est donnée.

Un CBF s'interprète ainsi : si une opération F_γ a comme arguments $expr_1 \dots expr_n$, ces arguments étant des éléments de A (donc des expressions graphiques) et que $expr_i$ appartient à l'ensemble des suites graphiques indexé (l'ensemble et non les suites qui lui appartiennent) par $indice_i$, alors le résultat produit par F_γ est une suite graphique qui appartient à l'ensemble des suites graphiques indexé par $indice_{résultat}$. Ce sont les CBF qui, en filtrant les éléments de A , vont spécifier la classe \underline{C} d'ensembles indexés (par des indices appartenant Δ). Dans chaque ensemble $C_\delta \in \underline{C}$ on aura les expressions graphiques élémentaires de X_δ et les expressions graphiques non élémentaires qui ont été spécifiées par un CBF avec $indice_{résultat} = \delta$. Par exemple, si l'on a $F_\gamma, X_{nom}, X_{art}$ du point (2) précédent et on a :

$$\Delta = \{\text{art, nom, sn}\}$$

$$\text{CBF-1} \quad \langle \langle F_\gamma, \langle \text{art, nom} \rangle, \langle \text{sn} \rangle \rangle \rangle$$

on aura :

$$\underline{C} = \{ C_{\text{art}}, C_{\text{nom}}, C_{\text{sn}} \}$$

$$C_{\text{art}} = \{\text{le}, \dots\}$$

$$C_{\text{nom}} = \{\text{garçon, livre}, \dots\}$$

$$C_{\text{sn}} = \{ (\text{le garçon})_\gamma, (\text{le livre})_\gamma \}$$

$$\cup_{\delta \in \Delta} C_\delta = \{\text{le, garçon, livre, } (\text{le garçon})_\gamma, (\text{le livre})_\gamma\}$$

Étant donné \underline{C} associé à un « langage », nous allons nommer $E \sim A^6$ l'ensemble $\cup_{\delta \in \Delta} C_\delta$. $E \sim A$ est en général, pour les LN, un sous-ensemble propre de A .

4 On sait que :

$$L'_X = \langle \langle L_X \sim A \rangle, R \rangle$$

et on sait qu'associés par les définitions à $L_X \sim A$ on a : $A, X_\delta \in \Delta, S$ (ensemble des CBF), Δ . Tous ces éléments sont les mêmes dans L_X et dans L'_X . En revanche, \underline{C} n'est pas associé à L'_X , alors que R va produire un nouvel ensemble associé à L'_X . Étant donné \underline{C} , R , utilisée comme une fonction sur C_δ , va donner Cat_δ , en effaçant les parenthèses des éléments de C_δ qui en ont (sur les éléments qui n'en ont pas, R produit les mêmes effets que $Id[\text{entité}]$). Ainsi, si l'on prend l'exemple du point précédent, on aura :

$$R(C_{\text{art}}) = Cat_{\text{art}} = \{\text{le}, \dots\}$$

$$R(C_{\text{nom}}) = Cat_{\text{nom}} = \{\text{garçon, livre}, \dots\}$$

$$R(C_{\text{sn}}) = Cat_{\text{sn}} = \{\text{le garçon, le livre}\}$$

À partir de \underline{C} et de R , on peut définir \underline{Cat} et l'ensemble EBF (Expressions Bien Formées⁷) :

$$\text{Étant donné } \underline{C} = \{C_1 - \delta, \dots, C_n - \delta\}$$

$$\underline{Cat} = \{Cat_1 - \delta', \dots, Cat_n - \delta'\} \text{ où } R(C_i - \delta) = Cat_i - \delta', \text{ avec } \delta = \delta'$$

$$EBF = \cup_{\delta \in \Delta} Cat_\delta = \{\text{le, garçon, livre, le garçon, le livre}\}$$

⁶ Cet ensemble n'a pas de nom consacré dans UG.

⁷ C'est l'ensemble « *meaningful expressions of L* » de UG.

C'est dans le cadre des points (1) à (4) ci-dessus que l'on doit se poser la question de la possibilité formelle de l'ambiguïté d'assignation catégorielle, dont on fait le rappel :
Est-ce que tout « langage non ambigu » va satisfaire la propriété d'unicité d'assignation catégorielle ? Ou, autrement dit, est-ce que l'objet $L \sim A \sim U$ est un objet formellement possible en UG ?

La réponse est que l'objet $L \sim A \sim U$ est un objet formellement possible en UG et que, par conséquent, tout « langage non ambigu » ne doit pas nécessairement satisfaire la propriété d'unicité d'assignation catégorielle.

Pour justifier la réponse nous indiquons comment obtenir dans un $L \sim A$ l'ambiguïté d'assignation catégorielle, c'est-à-dire pour faire en sorte qu'un $L \sim A$ soit $L \sim A \sim U$.

En ce qui concerne les expressions graphiques élémentaires, pour échapper à l'unicité d'assignation catégorielle, on doit incorporer la même expression graphique élémentaire à deux ensembles indexés différemment. Soit comme illustrations les exemples suivants :

- i $X_{art} = \{le, \dots\}$
 $X_{prclit} = \{le, \dots\} [pr[onom] clit [ique]]$
- ii $X_{nom} = \{juge, \dots\}$
 $X_{verbe} = \{juge, \dots\}$

Comme par définition $X\delta \subset C\delta$, on aura :

- $C_{art} = \{le, \dots\}$
- $C_{prclit} = \{le, \dots\}$
- $C_{nom} = \{juge, \dots\}$
- $C_{verbe} = \{juge, \dots\}$

et par conséquent l'unicité d'assignation catégorielle n'est pas satisfaite. Les restrictions sur l'ambiguïté d'un $L \sim A$ (cf. Section 2.1 points (ii) et (iii) sur le quintuplet qui définit un « langage ») ne font aucun obstacle à cette situation.

En ce qui concerne les expressions graphiques non élémentaires, il y a plusieurs procédés pour les obtenir, dont les deux suivants :

- i Spécifier deux CBF avec la même opération, avec les mêmes indices catégoriels comme arguments et donnant comme résultat des indices catégoriels différents.
- ii Spécifier deux CBF avec la même opération, avec des indices catégoriels différents comme arguments mais associés à des ensembles indexés qui ne satisfont pas l'unicité d'assignation catégorielle et donnant comme résultat des indices catégoriels différents.

Le procédé (i) est illustré par l'exemple factice suivant; $F\gamma$ toujours comme dans les exemples précédents :

- $CBF_i = \langle F\gamma, \langle \delta, \delta' \rangle \alpha \rangle$
- $CBF_j = \langle F\gamma, \langle \delta, \delta' \rangle \beta \rangle$

En supposant :

$$X_{\delta} = \{a, \dots\}$$

$$X_{\delta'} = \{b, \dots\}$$

on obtient :

$$C_{\alpha} = \{(a b)_{\gamma}, \dots\}$$

$$C_{\beta} = \{(a b)_{\gamma}, \dots\}$$

et par conséquent l'unicité d'assignation catégorielle n'est pas satisfaite. Les restrictions sur l'ambiguïté d'un $L \sim A$ ne font aucun obstacle à cette situation.

Le procédé (ii) est illustré par l'exemple non factice suivant; F_{γ} toujours comme dans les exemples précédents, et en plus C_{art} , C_{prclit} , C_{nom} , C_{verbe} comme ci-dessus. On a maintenant comme CBFs :

$$CBF_i = \langle F_{\gamma}, \langle art, nom \rangle sn \rangle$$

$$CBF_j = \langle F_{\gamma}, \langle prclit, verbe \rangle, sv \rangle$$

On obtient ainsi:

$$C_{sn} = \{(le\ juge)_{\gamma}, \dots\}$$

$$C_{sv} = \{(le\ juge)_{\gamma}, \dots\}$$

et par conséquent l'unicité d'assignation catégorielle n'est pas satisfaite. Les restrictions sur l'ambiguïté d'un $L \sim A$ ne font aucun obstacle à cette situation. Il s'agit d'une situation parfaitement réaliste et illustrée par les cas d'école :

La belle ferme la voile

gloses : *il y a une belle qui ferme la voile*

il y a une belle ferme qui voile quelque chose

Jacques le juge

gloses : *le juge Jacques*

Jacques juge quelqu'un

Nous allons noter $*expr$ une expression graphique, élémentaire ou non, telle que étant donné \underline{C}

$$\bullet *expr \in C_{\delta}, *expr \in C_{\delta'}; C_{\delta}, C_{\delta'} \in \underline{C}; \delta \neq \delta'$$

$$\bullet \text{pour } C_{\varepsilon} \in \underline{C}; \varepsilon \neq \delta \neq \delta', \text{ si } expr \in C_{\varepsilon}, \text{ alors } R(expr) \neq *expr$$

Une $*expr$ est ainsi une expression graphique qui appartient à deux ensembles indexés différents de \underline{C} et il n'existe pas une autre $expr$ appartenant à un autre ensemble indexé qui soit identique à $*expr$ par effacement des parenthèses.

2.2 Répercussions systémiques

L'analyse précédente montre que les « langages » $L \sim A \sim U$ existent bien dans le cadre des définitions portant sur les « langages non ambigus » de UG. La question est maintenant d'évaluer ce que cela entraîne pour le système d'ensemble et notamment en ce qui concerne :

- la définition de « langage » *syntactiquement ambigu* ;
- la définition d'expression *sémantiquement ambiguë* ;
- la notion de *traduction ambiguë*.

Comme il a déjà été dit, les deux premières définitions sont explicites dans UG et relèvent respectivement de UG p. 226 (Section 2 « Syntax ») et de UG p. 227 (Section 3 « Semantics : Theory of Meaning »). La notion de *traduction ambiguë* n'a pas été explicitée dans UG, mais elle est extrapolable à partir de la deuxième. On abordera ces trois points dans les trois sections qui suivent.

2.2.1 L'ambiguïté syntaxique

Étant donné :

$$L'_X = \langle \langle L_X \sim A \rangle, R \rangle$$

la clef de voûte pour définir formellement la notion d'ambiguïté syntaxique est la comparaison de l'ensemble $E \sim A_X$ associé à $L_X \sim A$, avec l'ensemble EBF_X associé à L'_X , ensembles définis, en suivant UG, dans la Section 2.1. En général, si :

- $\text{expr}', \text{expr}'' \in E \sim A$
- $R(\text{expr}') = \text{expr}, R(\text{expr}'') = \text{expr}; \text{expr} \in EBF$

alors expr est syntaxiquement ambiguë dans L'_X et le « langage » L'_X est syntaxiquement ambigu.

C'est donc l'existence de plus d'une expr dans $E \sim A$ associées par R à une seule et même expr dans EBF qui va faire considérer cette expr dans EBF comme syntaxiquement ambiguë et qui, par ricochet, permettra de déclarer ambigu le « langage » auquel EBF est associé.

Comme illustrations, nous reprenons $CBF-I$ et F_γ de l'exemple présenté dans la Section 2.1 (point 3) et supposons que \underline{C} satisfait la propriété d'unicité d'assignation catégorielle avec:

$$\begin{aligned}\underline{C} &= \{C_{\text{Sprep}}, C_{\text{nom}}, C_{\text{sn}}, C_{\text{art}}\} \\ C_{\text{Sprep}} &= \{((\text{ avec (des lunettes) }_\gamma)_\gamma, \dots)\} \\ C_{\text{nom}} &= \{\text{fille}, \dots\} \\ C_{\text{art}} &= \{\text{la}, \dots\} \\ C_{\text{sv}} &= \{\text{regarde}, \dots\}\end{aligned}$$

Par ailleurs, nous ajoutons :

- CBF-2 $\langle \langle F_\gamma, \langle n, \text{sprep} \rangle, n \rangle$
CBF-3 $\langle \langle F_\gamma, \langle sv, sn \rangle, sv \rangle$
CBF-4 $\langle \langle F_\gamma, \langle sv, \text{sprep} \rangle, sv \rangle$

Les nouveaux CBF vont permettre de spécifier les expressions graphiques suivantes, qui sont des éléments du $C_{\langle \text{indice catégoriel} \rangle}$ indiqué entre crochets à leur droite :

- i (fille (avec (des lunettes) γ) γ) γ [C_{nom}]
ii (regarde (la (fille (avec (des lunettes) γ) γ) γ) γ) γ [C_{sv}]
iii (regarde (la (fille) γ) γ) γ [C_{sv}]
iv (regarde (la (fille) γ) γ (avec (des lunettes) γ) γ) γ [C_{sv}]

Les expressions graphiques (ii), (iii) et (iv) sont des éléments de C_{sv} et, en tant que tels, elles sont différentes deux à deux; cette différence étant trouvée à partir de la définition générale suivante⁸ :

Étant donné les expressions graphiques a et b suivantes :

a $x_1 x_2 \dots x_n$

b $y_1 y_2 \dots y_n$

l'expression graphique a est différente de l'expression graphique b , si pour $\langle x_i, y_i \rangle$, $x_i \neq y_i$.

Si l'on se tient aux expressions graphiques (ii) et (iv), on a le C_{sv} suivant, et avec R on obtient Cat_{sv} :

$C_{sv} = \{ (\text{regarde} (\text{la} (\text{fille} (\text{avec} (\text{des lunettes})\gamma)\gamma)\gamma)\gamma, \\ (\text{regarde} (\text{la} (\text{fille})\gamma)\gamma (\text{avec} (\text{des lunettes})\gamma)\gamma)\gamma, \dots \}$
 $Cat_{sv} = \{ \text{regarde la fille avec des lunettes}, \dots \}$

De telle manière qu'étant donné :

$L'y = \langle \langle L_y \sim AU \rangle, R \rangle$

avec :

$E \sim A_y = \{ (\text{regarde} (\text{la} (\text{fille} (\text{avec} (\text{des lunettes})\gamma)\gamma)\gamma)\gamma, \\ (\text{regarde} (\text{la} (\text{fille})\gamma)\gamma (\text{avec} (\text{des lunettes})\gamma)\gamma)\gamma, \dots \}$
 $EBF_y = \{ \text{regarde la fille avec des lunettes}, \dots \}$

on doit conclure, selon les définitions de UG, que *regarde la fille avec des lunettes* est une expression graphique syntaxiquement ambiguë dans $L'y$, et que $L'y$ est syntaxiquement ambigu. Et c'est en effet la solution intuitive souhaitée pour pouvoir rendre compte des gloses observationnelles suivantes :

- (regarde (la (fille (avec (des lunettes) γ) γ) γ) γ) γ
glose : *regarde la fille qui possède des lunettes*

⁸ Cette définition de différence n'est pas donnée en UG; elle vient de l'utilisation faite ici de la notation traditionnelle par parenthèses indexées.

- (regarde (la (fille) γ) γ (avec (des lunettes) γ) γ) γ
glose : *regarde la fille en utilisant des lunettes*

La question se pose donc de déterminer, étant donné cette fois

$$L'_Z = \langle \langle L_Z \sim A \sim U \rangle, R \rangle \text{ et } \underline{C}_Z \text{ associé}$$

quel est le statut d'une $*expr_i$, tel que $R(*expr) = *expr_r$ (rappel : $*expr$ appartient à deux ensembles différemment indexés de \underline{C}_Z et il n'existe pas une autre $expr$ appartenant à un autre ensemble indexé qui soit identique à $*expr$ par effacement des parenthèses, cf. Section 2.1 *in fine*).

D'après les définitions de UG, il faut conclure que $*expr_r$ n'est pas une expression syntaxiquement ambiguë dans L'_Z , et l'existence des expressions graphiques de type $*expr_r$ ne rend pas « langage ambigu » un « langage » tel que L'_Z . Et cela, en raison du fait exposé dans le paragraphe suivant.

L'ambiguïté syntaxique d'une expression graphique se définit en comparant les ensembles $E \sim A$ et EBF et en appliquant R sur les éléments de EBF . Pour qu'une $expr_k \in EBF$ soit considérée syntaxiquement ambiguë, il faut au moins deux expressions telles que $expr_i$, $expr_j \in E \sim A$, et $expr_i \neq expr_j$. Or $*expr$, par définition, est une seule expression dans $E \sim A_Z$, bien que relevant, du fait même qu'il s'agit d'un « langage non ambigu » qui ne satisfait pas l'unicité d'assignation catégorielle, d'au moins deux ensembles indexés différents de \underline{C}_Z .

Si nous revenons au dernier exemple de la Section 2.1, où on avait obtenu

$$C_{sn} = \{ ((\text{le juge}) \gamma \dots) \}$$

$$C_{sv} = \{ ((\text{le juge}) \gamma \dots) \}$$

et où l'unicité d'assignation catégorielle n'est pas satisfaite, on a comme résultat que l'expression *le juge* n'est pas expression syntaxiquement ambiguë dans EBF , même s'il y a une expression $(\text{le juge}) \gamma$ assignée à l'ensemble indexé des syntagmes nominaux et une expression $(\text{le juge}) \gamma$ assignée à l'ensemble indexé des syntagmes verbaux, car on a :

$$E \sim A = \{ ((\text{le juge}) \gamma \dots) \}$$

$$EBF = \{ \text{le juge}, \dots \}$$

et il n'y a pas $expr$ telle que $R(expr) = \text{le juge}$, $expr \neq (\text{le juge}) \gamma$, $expr \in E \sim A$.

La même analyse et pour les mêmes raisons doit être faite pour les expressions graphiques élémentaires : les expressions *juge*, *le* ne sont pas syntaxiquement ambiguës dans l'exemple.

L'unicité d'assignation catégorielle n'est donc pas une exigence formelle sur les « langages non ambigus », mais la non-unicité d'assignation catégorielle ne permet pas d'exprimer une ambiguïté syntaxique.

2.2.2 L'ambiguïté sémantique

Étant donné :

$$L'_X = \langle \langle L_X \sim A \rangle, R \rangle$$

l'assignation de signification à L'_X (« meaning assignment for » L'_X , UG p. 227) se définit pour ainsi dire en deux étapes, la première précisant ce qu'est une *interprétation* (anglais *interpretation*) de L'_X , et la deuxième définissant l'homomorphisme unique d'assignation de signification. C'est dans le cadre de cet homomorphisme unique d'assignation de signification que la notion formelle de *means* est définie et avec elle la notion d'expression graphique *sémantiquement ambiguë*. Ce sont ces définitions qui seront par la suite révisées à partir de

$$L'_Z = \langle \langle L_Z \sim A \sim U \rangle, R \rangle$$

avec donc une ou plusieurs expressions de type **expr* appartenant à EBF_Z . Nous synthétisons par la suite ce qui va permettre d'évaluer le sort formel de ces **expr* en présentant les éléments clefs des définitions par rapport à $L'_X = \langle \langle L_X \sim A \rangle, R \rangle$.

- 1 Une « interprétation » est un triplet $\langle B, G_\gamma, f \rangle$ où
 - B est un ensemble d'objets quelconques nommés *meanings* (que nous traduisons par « significations »; sur la nature arbitraire de ces objets, cf. EGM Section 3.2.1 Ob. 1 *in fine*);
 - G_γ est un ensemble d'opérations;
 - f est une fonction (sur laquelle on revient ci-dessous).
- 2 Étant donné $\langle A, F_\gamma \rangle$ dans le quintuplet qui définit $L_X \sim A$ et $\langle B, G_\gamma \rangle$ de l'« interprétation », $\langle A, F_\gamma \rangle$ et $\langle B, G_\gamma \rangle$ sont des algèbres similaires (c'est-à-dire qu'il existe une fonction bijective entre les ensembles F_γ et G_γ , et les opérations associées par cette fonction ont la même arité).
- 3 La fonction f a comme domaine $\cup_{\delta \in \Delta} X_\delta$ associé à $L_X \sim A$. La fonction g a comme domaine $\cup_{\delta \in \Delta} C_\delta$ (c'est-à-dire l'ensemble $E \sim A$; on sait par ailleurs que par définition $\cup_{\delta \in \Delta} X_\delta \subset E \sim A$). Par ailleurs $f \subset g$, ce qui signifie que si $expr \in X_\delta$, $f(expr) = b$, alors pour $expr \in C_\delta$, $g(expr) = b$.

C'est dans le cadre planté par (1), (2) et (3) ci-dessus que UG définit la notion de *signifie* (anglais *means*) et la notion d'expression sémantiquement ambiguë dans un « langage » (« semantically ambiguous » dans L'_X), celle-ci d'une manière parallèle à la définition d'expression syntaxiquement ambiguë, révisée dans la Section précédente.

Les définitions de ces notions peuvent s'exprimer comme suit.

Étant donné

$$L'_X = \langle \langle L_X \sim A \rangle, R \rangle$$

- $\text{expr} \in \text{EBF}$ signifie b dans L'_X si
 $\text{expr} \in \text{EBF}$, $\text{expr}' \in E \sim A$, $R(\text{expr}') = \text{expr}$, $g(\text{expr}') = b$,
- $\text{expr} \in \text{EBF}$ est sémantiquement ambiguë dans L'_X si
 $\text{expr}' \in E \sim A$, $\text{expr}'' \in E \sim A$, $R(\text{expr}') = \text{expr}$, $R(\text{expr}'') = \text{expr}$
 $g(\text{expr}') = b$, $g(\text{expr}'') = b'$; $b, b' \in B$

La clef pour déclarer *sémantiquement ambiguë* une *expr* appartenant à *EBF* est qu'elle soit associée par *R* à deux expressions nécessairement différentes de $E \sim A$. Il est donc clair qu'une *expr* ne peut être sémantiquement ambiguë que si elle est syntaxiquement ambiguë.

Nous savons déjà qu'une **expr* ne peut être syntaxiquement ambiguë. Une **expr* ne peut donc pas non plus être sémantiquement ambiguë. Si nous revenons à l'exemple précédent, la suite graphique *le juge* n'est ni syntaxiquement ambiguë ni sémantiquement ambiguë.

Il nous reste à nous poser la question : quel est le statut d'une **expr* par rapport à l'homomorphisme unique d'assignation de signification? Ou autrement dit, qu'est-ce qui se passe avec une **expr* par rapport à l'« interprétation »?

Notre position est qu'il ne se passe rien : une **expr* est un non-sens pour le système d'ensemble, si on pense que ce système est un objet qui, par son formalisme à lui, détermine les résultats. Or étant donné que l'on est dans une situation où

$$R(\text{expr}) = * \text{expr}, \text{expr} \in C\delta \text{ expr} \in C\delta', \delta \neq \delta'$$

il est faux de dire que le système *impose* une seule interprétation à **expr*⁹ : le système est incapable de définir en entrée d'une de ses fonctions clés (l'homomorphisme unique d'assignation de signification) des objets différenciés pour obtenir des résultats différents. C'est à l'utilisateur du système qu'il revient de décider arbitrairement si à **expr* on doit assigner un « meaning » selon $C\delta$ ou selon $C\delta'$, ce qui est un non-sens formel.

Ce non-sens arrive grâce à une erreur profonde de modélisation de l'objet langue naturelle. Dans les langages logiques ce sont les parenthèses qui lèvent les ambiguïtés et, étant donné la commodité de différencier les entités élémentaires par leurs caractéristiques graphiques, on a soigneusement évité dans ces langages de ne pas satisfaire la propriété d'unicité d'assignation catégorielle. Cela n'a pas de sens intuitif de parler d'un langage logique *désambiguïsé* : les langages logiques naissent non ambigus. Les langues naturelles se propagent dans le temps et dans l'espace comme des langages constitutivement ambigus. Si on considère ces objets comme des objets à étudier, certes comme langages, mais avec

⁹ Cela contrairement à ce qui est dit dans [Dowty & al. 81]; cf. document ci-avant EGM Section 1.1.

les exigences d'une science empirique, on s'aperçoit qu'une de leurs caractéristiques est que les expressions élémentaires qu'on doit leur attribuer sont perceptivement ambiguës – ce ne sont pas leurs caractéristiques sensibles qui peuvent distinguer les classes dans lesquelles on va les distribuer pour étudier leur combinatoire (cf. dans la Section 7.2 du document EGM ci-avant la formulation VI). Et on ne peut pas désambiguïser les langues naturelles en utilisant des parenthèses – même indexées – résultant des opérations, comme la modélisation de UG a essayé de le faire, en important en linguistique la tradition conceptuelle et notationnelle de la logique : c'est un point clef sur lequel la naturalité de la langue naturelle est à l'œuvre.

2.2.3 L'ambiguïté de traduction

UG n'introduit aucune notion relative à l'ambiguïté de la traduction dans la Section 5 consacrée à la « théorie » de la traduction (« Theory of translation », p. 231-233), mais on peut appliquer quasiment à la lettre ce qui vient d'être dit par rapport à l'assignation de significations et à l'ambiguïté sémantique. Le cadre où se définit la traduction peut être fixé par les trois points suivants.

1 Étant donné :

$$L'_X = \langle \langle L_X \sim A \rangle, R \rangle$$

la traduction s'effectue de L'_X sur LI^{10} et, plus exactement, sur Ad^{LI} (cf. immédiatement ci-dessous). À chacun de ces langages on suppose associées les entités classiques : $\underline{C}_X, \Delta_X, E \sim A_X$ à L_X ; EBF à L'_X . Associé à LI on aura Ad^{LI} : algèbre dérivée de LI qui résulte de l'application des opérations polynomiales sur l'algèbre $\langle A, F_\gamma \rangle_{LI}$ du quintuplet définissant LI comme langage. $\langle A, F_\gamma \rangle_X$ (résultant du quintuplet définissant L_X) et Ad^{LI} sont des algèbres similaires, les deux étant associées à Γ (ensemble des indices des opérations). Les opérations polynomiales permettant d'obtenir Ad^{LI} seront notées H_γ avec $\gamma \in \Gamma$. Les entités $\underline{C}, \Delta, E \sim A$ sont associées au quintuplet Ad^{LI} (cela est important pour $E \sim A$) mais on les notera pour simplifier l'écriture avec LI comme indice : on aura ainsi $\underline{C}_{LI}, \Delta_{LI}, E \sim A_{LI}$; de même on continuera à parler de traduction de L'_X sur LI .

2 Une Base de traduction de L'_X sur LI est un triplet $\langle g', H_\gamma j \rangle_{\gamma \in \Gamma}$ dans lequel (sur H_γ cf. le point précédent) :

- g' est une fonction de Δ_X sur Δ_{LI}
- j est une fonction de domaine $\cup_{\delta \in \Delta_X} X\delta$ et codomaine $C\delta \in \underline{C}_{LI}$ telle que si $expr \in X\delta, \delta \in \Delta_X, j(expr) = expr', expr' \in C\delta',$ et $g'(\delta) = \delta'.$

3 La fonction k est l'homomorphisme unique de $\langle A, F_\gamma \rangle_X$ sur Ad^{LI} , et $j \subset k$.

¹⁰ Rappel : nous avons simplifié le problème en nous centrant sur LI et non sur un "langage" arbitraire quelconque.

C'est dans le cadre des points (1), (2) et (3) précédents que la notion « $expr_{LI}$ est une traduction de $expr_{L'x}$ » est définie. On peut l'exprimer ainsi :

Étant donné :

$$L'_x = \langle \langle L_x \sim A \rangle, R \rangle \text{ et } LI$$

$expr_{LI} \in E \sim A_{LI}$ est une traduction de $expr \in L_x$ si

$$expr \in EBF_{L'_x}, expr' \in E \sim A_{L_x}, R(expr') = expr, k(expr') = expr_{LI}$$

En calquant la définition de traduction ambiguë sur celle de sémantiquement ambigu, on peut ajouter (notion non explicitée dans UG) :

$expr_{LI'}, expr_{LI''} \in E \sim A_{LI}$ sont des traductions ambiguës de $expr \in L_x$ si

$$expr \in EBF_{L'_x}, expr' \in E \sim A_{L_x}, R(expr') = expr, k(expr') = expr_{LI'}$$

$$expr'' \in E \sim A_{L_x}, R(expr'') = expr, k(expr'') = expr_{LI''}$$

Pour les mêmes raisons que celles signalées dans la Section précédente, à une $*expr$ de L'_x ne peuvent pas être associées des traductions ambiguës et, pour la fonction de traduction (ambiguë ou pas), les $*expr$ sont un non-sens.

3 Ambiguïté d'assignation catégorielle dans PTQ

Nous synthétisons par la suite les divergences et les points de contact entre UG et PTQ, ce qui devrait permettre de poser dans le cadre de PTQ les deux questions sur l'ambiguïté.

1 Dans PTQ il n'existe ni utilisation explicite du quintuplet qui définit un « langage », ni utilisation d'un « langage non ambigu » ni relation R ni a fortiori contraintes assurant la non-ambiguïté. Mais les éléments suivants du quintuplet définissant un « langage » sont utilisés :

- L'ensemble d'indices catégoriels (maintenant malheureusement nommés *catégories*) est conservé ; Δ de UG est, dans PTQ, noté Cat . La notation $\cup_{\delta \in \Delta} X_{\delta}$ de UG est devenue $\cup_{A \in Cat} B_A$ de PTQ, X_{δ} de PTQ se note B_A (B de *basic expression*), mais les objets dénotés par $\cup_{\delta \in \Delta} X_{\delta}$ et par $\cup_{A \in Cat} B_A$, par X_{δ} et par B_A ont les mêmes caractéristiques formelles.
- Les opérations notées F_{γ} de UG subsistent, sans pourtant que Γ (tel que $\gamma \in \Gamma$) soit explicitement reconnu; en revanche on note dans PTQ F_n avec $n \leq 0$ sans préciser par là si l'on veut un ensemble infini dénombrable d'opérations. Comme il n'y a pas de contraintes assurant la non-ambiguïté, on ne doit pas supposer l'utilisation des parenthèses insérées par chaque opération et indexées par rapport à cette opération.
- Les CBF subsistent et sont maintenant labellisées *règles*. On ne voit rien de différent entre l'ensemble S des CBFs dans UG et un ensemble de « règles » de PTQ; d'ailleurs les « règles » de PTQ son labellisées $S_1 \dots S_n$. Mais, et cela est important pour le problème qui nous occupe, il n'y a pas d'ensemble $E \sim A$ dans PTQ – ce qui est normal, étant donné l'absence de R et des stipulations sur la non-ambiguïté – et il n'y a pas non plus d'ensemble EBF défini comme l'union des expressions d'une famille d'ensembles.

Pourtant en PTQ on continue à parler d'expression bien formée (*meaningful expression*); il s'agit d'une expression graphique d'un quelconque des ensembles indexés d'expressions, élémentaires ou non, et dont les indices sont des éléments de *Cat* (dans la notation PTQ). On note les ensembles indexés d'expressions bien formées par P_A , avec $A \in \text{Cat}$ (et P mnémotechnique de « phrase »); on sait que $B_A \subset P_A$.

2 Il n'y a pas dans PTQ, contrairement à UG, de définition explicite sur l'ambiguïté, qu'elle soit syntaxique ou sémantique.

On sait que PTQ se donne un objet – l'arbre dérivationnel (*analysis tree*) et qu'il utilise des règles de traduction définissant la relation de traduction (*translation relation*) et qu'il distingue la possibilité d'avoir « infinitely many analysis trees », lesquels « will differ only inessentially » des situations où une expression « has two essentially different analyses ». Ces points sont ceux qui concernent les répercussions systémiques de l'ambiguïté d'assignation catégorielle; on les reprend dans la Section 3.2, mais avant, dans la Section 3.1, on aborde la question de la possibilité formelle, plus directement concernée par le premier des points que l'on vient d'évoquer.

3.1 Possibilité formelle

Il est clair que, s'il n'existe pas de contraintes assurant la non-ambiguïté, il n'y aura pas d'empêchement potentiel à insérer la même expression graphique élémentaire dans des ensembles d'expressions indexés par des indices catégoriels différents. Les objets dans $\cup_{A \in \text{Cat}} B_A$ de PTQ fonctionnent comme ceux de $\cup_{\delta \in \Delta X \delta}$ de UG, et rien n'empêche formellement l'appartenance d'une même expression à deux ensembles indexés différemment.

Même raisonnement par rapport aux résultats susceptibles d'être obtenus par les CBF. Par les mêmes deux procédés décrits pour UG (cf. Section 2.1), on doit pouvoir associer une même expression graphique non élémentaire à deux ensembles indexés différemment.

Tout cela si l'on reste à l'analyse des formulations. Curieusement ou non, dans les descriptions concrètes, on n'a jamais oublié, ni dans PTQ ni dans les ouvrages qui l'ont présenté, de respecter l'unicité d'assignation catégorielle.

3.2 Répercussions systémiques

Il est difficile de se faire une idée précise de ce qu'est un arbre de dérivation. L'objet est présenté par un exemple dans [PTQ p. 254], où on doit se contenter d'un renvoi à EFL et des explications suivantes :

To each node we attach a meaningful expression, together, in case that expression is not basic, with the index of that structural operation [...] which we understand as having been applied in obtaining the expression in question; the nodes dominated by any node are to be occupied by the

expressions to which the structural operation is understood as having been applied in obtaining the expression occupying the superior node.

Et PTQ ajoute quelques lignes plus loin:

A precise characterization of an *analysis tree* in the sense of these remarks would be routine and will not be given here [suit un renvoi à EFL].

C'est immédiatement après cette présentation court-circuitée que l'on introduit la différence entre l'existence inessentielle d'une pluralité d'arbres lesquels « will all lead to the same semantical results » (PTQ p. 255), et la pluralité d'arbres qui expriment « genuinely (that is, semantically) ambiguous sentences » (PTQ p. 255). Mais il serait vain de chercher à trouver dans PTQ une caractérisation formelle de cette différence.

Il serait tout aussi vain de chercher l'explicitation du lien formel entre présentations à la UG et présentations à la PTQ. Sur ce point crucial [PTQ p. 254] est encore elliptique :

If it were desired to construct a corresponding [avec les arbres dérivationnels] unambiguous language, it would be convenient to take the analysis trees themselves as the expressions of that language; it would then be obvious how to characterize (in keeping with [UG]) the structural operations of that language and the corresponding relation between its expressions and those of ordinary English. For present purposes, however, no such construction is necessary¹¹.

Le problème de fond que nous voyons avec les arbres dérivationnels est que ces arbres sont essentiellement une procédure, une manière de faire. Ainsi PTQ introduit l'exemple d'arbre dérivationnel qui sert à définir cet objet en disant (c'est nous qui soulignons en gras) :

We may indicate **the way** in which this sentence has just been constructed by means of the following *analysiss tree* (p. 254).

et la question est de savoir quelle est l'information que cette manière de faire construit au fur et à mesure et qu'elle laisse dans les nœuds de l'arbre; comme on le verra par la suite, ce point est crucial pour le problème qui nous occupe des répercussions systémiques du non-respect de l'unicité d'assignation catégorielle ; et tout le monde n'a pas interprété les nœuds de l'arbre de la même manière.

Sans bien savoir ce qu'est formellement un arbre dérivationnel, il est difficile de faire le lien formel de l'ensemble des arbres spécifiés par une description avec la définition de l'ensemble des CBF et les ensembles P_A (ensembles indexés des suites bien formées) spécifiés par cette même description. Sur ce dernier point – le lien entre CBF et les ensembles P_A –, PTQ continue à être elliptique. Il suggère que les CBF constituent « a simultaneous inductive definition of the sets P_A », il ajoute que puisque « however,

¹¹ La note de bas de page ajoutée à ce point en PTQ est révélatrice des difficultés rencontrées lorsqu'il s'agit de procéder à des descriptions linguistiques en utilisant des "langages non ambigus". La note (5) de PTQ dit ceci : "This way of constructing an underlying unambiguous language, though convenient here, would be unsuitable in connection with fragments of natural language exhibiting greater syntactical complexities of certain sorts."

inductive definitions of this form are somewhat unusual, it is perhaps in order to state a corresponding explicit definition ». Or, la définition proposée, la voici :

...the sets P_A (for $A \in \text{Cat}$) are the smallest sets satisfying S1-S17 [C'est-à-dire les CPF de PTQ]; that is to say, $\langle P_A \rangle_{A \in \text{Cat}}$ is the unique family of sets indexed by Cat such that (1) $\langle P_A \rangle_{A \in \text{Cat}}$ satisfies S1-S17, and (2) whenever $\langle P'_A \rangle_{A \in \text{Cat}}$ is a family of sets indexed by Cat , if $\langle P'_A \rangle_{A \in \text{Cat}}$ satisfies S1-S17, then $P_A \subseteq P'_A$.

Et comme démonstration, on a le commentaire suivant :

(It is easily shown, using an idea I believe to have originated with Dr. Perry Smith, that there is exactly one family of sets satisfying these conditions) (PTQ p. 253).

Nous avons ainsi trois objets formels distincts, qui ont certainement des liens formels entre eux, mais que nous ne connaissons pas :

- ensemble EBF défini en UG à partir de $E \sim A$;
- ensembles indexés de PTQ qui devraient résulter de la définition inductive suggérée dans PTQ ;
- ensembles d'arbres dérivationnels tels qu'on peut les imaginer à partir de leur présentation rapide et des exemples qui sont donnés, arbres que l'on peut conceptuellement imaginer comme une manière quasi algorithmique d'analyser une expression graphique non élémentaire.

Un bon indice que les liens entre ces objets ne doivent pas être évidents, est le fait que les arbres dérivationnels ont été interprétés autrement que comme il est dit dans UG et pratiqué dans la notation des exemples qui s'y trouvent.

Ainsi dans [Chambreuil 89 p. 57-58] on a explicitement introduit dans l'explicitation formelle de la notion de suite dérivationnelle « qui peut également recouvrir la notion de graphe adoptée pour la présentation de PTQ » une modification par rapport à UG. On y lit :

Il suffit de remplacer [dans la définition donnée pour UG] la condition b de $c1$ par (b') à $a_1 \dots a_{n\gamma}$ sont respectivement associées les indices $\delta_1 \dots \delta_n$ [...]

alors que la condition b de $c1$ était :

$$(b) \langle a_1 \dots a_{n\gamma} \rangle \in C\delta_1 x \dots x C\delta_n \gamma$$

La différence par rapport à la présentation de PTQ consiste dans ceci. D'après [Chambreuil 89 p. 57-58] ce sont les expressions qui deviennent indexées : si, dans le graphe qui aurait pu être donné pour UG, on exige qu'une expression appartienne à un ensemble qui, lui, est indexé – et non l'expression –, par contre dans le graphe de PTQ, selon Chambreuil, c'est l'expression qui le devient. Ceci autorise à dire (c'est nous qui soulignons en gras) :

Aussi, nous dirons également que le dernier élément d'un graphe dérivationnel est une expression significative; **ce sera en fait un assemblage auquel les règles permettent d'attribuer un indice catégoriel** [Chambreuil 89 p. 58].

L'autre différence entre *arbre à la PTQ façon Montague* et *arbre à la PTQ façon Chambreuil* vient du fait que le premier enregistre dans le nœud l'opération qui l'a produit et, comme conséquence du point précédent, il n'enregistre pas l'indice catégoriel, tandis que le deuxième enregistre le CBF (c'est-à-dire la « règle »), et comme conséquence du point précédent, il associe l'indice catégoriel à l'expression obtenue. Ainsi pour *John seeks a unicorn* le premier dira que dans les deux dernières lignes de l'arbre on aura (i) alors que le deuxième aura (ii).

- i John seeks a unicorn, F₄
 John seek a unicorn, F₅
- ii John seeks a unicorn (t), S₄
 John (T) seek a unicorn (IV), S₅

Tout cela a des conséquences directes sur notre problème consistant à évaluer les répercussions systémiques du non-respect dans PTQ de l'unicité d'assignation catégorielle. Elles peuvent se synthétiser ainsi : la notion d'ambiguïté est associée dans PTQ à la traduction des arbres; les répercussions systémiques se verront donc au niveau de la traduction des arbres, objets qui eux-mêmes ne sont pas formellement clairs. C'est donc sans surprise que l'on arrivera à des conclusions qui seront conditionnées par les idées que l'on se fait des arbres.

Dans PTQ – contrairement à UG – on ne parle plus d'homomorphisme unique d'assignation de signification, on n'a pas une base de traduction, et on n'a pas une « translation function » qui soit un homomorphisme unique entre deux algèbres (celle du « langage non ambigu » et celle du « langage » cible, en l'occurrence l'algèbre dérivée de LI). En PTQ, la traduction se fait par :

- la fonction f qui traduit les indices catégoriels [nommés, rappelons-le *catégories* et obtenus par définition récursive] de Cat [équivalent en PTQ du Δ de UG] sur les types de LI ;
- la fonction g qui traduit un sous-ensemble des expressions graphiques élémentaires ;
- les règles de traduction.

La fonction f , telle qu'elle est formulée dans PTQ, traduisant de manière générale les éléments qui ont permis la définition récursive des indices catégoriels, est inexistante en UG. Elle ne pose pas de problème d'interprétation; nous la prenons telle qu'elle a été formulée.

Le domaine de la fonction g peut se caractériser ainsi. Soit l'ensemble d'expressions graphiques élémentaires (notées façon PTQ) $\cup_{A \in Cat} B_A$. Nous allons dire :

$$\bigcup_{A \in \text{Cat}} B_A = \text{EL}^g \cup \text{EL}^r, \text{ avec } \text{EL}^g \cap \text{EL}^r = \{\}$$

Dans $EL\mathcal{G}$ nous avons les expressions graphiques élémentaires qui sont traduites par g ; dans EL' celles qui sont traduites par une règle; l'ensemble des expressions graphiques se

partitionne en $EL\mathcal{S}$ et EL' . Cette exigence de partition, qui fait qu'une même *expr* ne peut pas appartenir à $EL\mathcal{S}$ et à EL' , même si elle n'est pas dite en toutes lettres dans PTQ, se dégage de [PTQ p. 260] où le domaine de g est défini. Par ailleurs, g est caractérisée (*ibidem*) comme une fonction (« fixed biunique fonction ») dont le codomaine sont des expressions de LI dont le type est exigé par f , de telle manière que si *expr* appartient à l'ensemble indexé par A , l'*expr'* de LI dans laquelle *expr* sera traduite sera de type $f(A)$.

Les règles de traduction sont classées en règles basiques (*Basic rules*) et les autres, celles-ci recevant en sous-classement des étiquettes variées qui ne nous intéressent pas.

Par rapport à la question de l'ambiguïté, [PTQ p. 263] affirme que :

the translation relation" [laquelle est définie par les "rules of translation" PTQ p. 261] is of course not a function; a meaningful expression of English may translate into several different expressions of intensional logic. We could, however, speak of *the translation* [souligné dans l'original] of a given meaningful expression of English corresponding to any given analysis tree for that expression; the rather obvious definition of this notion will be omitted here.

Pour ce qui est de l'affirmation « a meaningful expression of English may translate into several different expressions of intensional logic », il faut l'examiner selon qu'il s'agisse d'expressions élémentaires, celles appartenant donc à $\cup_{A \in Cat} B_A$, et celles qui ne sont pas élémentaires et qui ont été assignées par un CBF à un ensemble indexé.

Étant donné que g est une fonction dont le domaine est $EL\mathcal{S}$, il est impossible qu'une *expr* appartenant à $EL\mathcal{S}$ soit traduite par deux objets différents dans LI. Par ailleurs $EL\mathcal{S}$, étant caractérisé comme le complément dans $\cup_{A \in Cat} B_A$ de l'ensemble noté ici EL' , celui-ci étant caractérisé (cf. [PTQ p. 260]) soit par énumération des expressions graphiques qui s'y trouvent (p. ex. *be*) soit par inclusion de l'ensemble indexé par un indice catégoriel (en PTQ, c'est l'ensemble indexé par T), il n'existe pas non plus la possibilité de traduire de manière différenciée une même *expr* qui aurait été assignée à un ensemble indexé inclus dans $EL\mathcal{S}$ et assignée aussi, comme c'est possible de le faire formellement (cf. Section 3.1), à un ensemble indexé différemment inclus dans EL' , car, en raison des exigences des règles de traduction, $EL\mathcal{S}$ et EL' sont disjoints.

Nous débouchons donc ici sur la même conclusion que pour UG : si par permissivité formelle on fait appartenir une même *expr* à deux ensembles indexés différemment les deux étant inclus dans $EL\mathcal{S}$, ou l'un étant inclus dans $EL\mathcal{S}$ et l'autre dans EL' on aboutit à un non-sens formel.

Supposons que chaque description à la PTQ, pour éviter ce non-sens formel, vienne accompagnée des recommandations ad hoc indiquant la catégorie à choisir comme point de départ de la fonction f , et/ou, si une *expr* appartient aussi bien à $EL\mathcal{S}$ qu'à EL' , indiquant à quoi on doit décider de la faire appartenir. Quel choix faire pour, p. ex., *mange* (ou *eats*) transitif et intransitif. On aurait :

Cat = {TV, IV,...}

$B_{TV} = \{ \text{mange (eats),...} \}$
 $B_{IV} = \{ \text{mange (eats),...} \}$
 $f(TV) = \langle \langle s, \langle \langle s, \langle \langle s, e \rangle, t \rangle \rangle, t \rangle \rangle, \langle \langle s, e \rangle, t \rangle \rangle$
 $f(IV) = \langle \langle s, e \rangle, t \rangle$
 $g(\text{mange (eats)}) = x, \text{ et } x \in Pf(TV)$
 $g(\text{mange (eats)}) = x', \text{ et } x' \in Pf(IV)$

Comme g est une fonction, $x = x'$. Si nous voulons que x soit une constante, comme les constantes de $Pf(TV)$ sont notées $c_n, f(TV)$ et celles de $Pf(IV)$ sont notées $c_m, f(IV)$, nous aurions besoin que $c_n, f(TV) = c_m, f(IV)$, ce qui est impossible, car par définition de f , $f(IV) \neq f(TV)$. Et si l'on choisit arbitrairement l'une de ces deux solutions, le résultat sera erroné pour les expressions qui réclament l'autre.

Le raisonnement serait analogue si l'on avait une expression graphique élémentaire appartenant à $EL\mathcal{S}$ et à EL' (cela dans le cas où on n'accepterait pas $EL\mathcal{S} \cap EL' = \{\}$). Dans cette situation on devrait indiquer dans une recommandation ad hoc si l'expression doit être traitée selon son appartenance à $EL\mathcal{S}$ ou à EL' . Ainsi, p. ex., *la* en français, en tant qu'appartenant à EL' , exige une traduction comme pour *the* en (i) ci-avant, alors qu'en tant que note de musique, comme (ii).

- i $\lambda P[\lambda Q \exists x [\forall x [P\{x\} \leftrightarrow x=y] \cap Q\{y\}]]$
- ii $c_n, f(N)$

Le choix arbitraire donnera de mauvais résultats si l'on choisit (i) pour traiter *il a donné le la*, et, si l'on choisit (ii), pour traiter *la musique est finie*.

Ces situations d'ambiguïté sont donc formellement intraitables par le non-respect de l'unicité d'assignation catégorielle des expressions graphiques élémentaires. Les situations évoquées ne sont pas l'exception dans les langues naturelles; elles ne sont pas non plus le fruit d'une observation malveillante des cas bizarres. Les situations effectives sont encore plus compliquées que les exemples précédents ne le laissent entendre. Les cas où une expression graphique élémentaire devrait être assignée à deux ensembles différemment indexés ne sont pas rares : le *la* du français n'en exige pas moins de trois; en anglais, le *ninety* de PTQ n'est traité que comme $c_n, f(N)$ alors qu'en anglais on a aussi *ninety boys*, *the ninety boys*, *the room ninety*, que l'on voit mal comment traiter par $c_n, f(N)$.

En ce qui concerne l'ambiguïté des expressions graphiques non élémentaires, on ne voit rien d'explicite en PTQ (cf. ci-dessus Section 3.1) qui interdise de les spécifier, et cela en utilisant les mêmes procédés que dans UG (cf. ci-dessus Section 2.1). Est-ce qu'on peut les traduire de manière ambiguë ? La réponse pour UG a été non; pour PTQ elle est : cela dépendra de ce que l'on croit être un arbre dérivationnel et des expressions graphiques elles-mêmes.

Comme on l'a déjà remarqué, l'objet *arbre de dérivation ou dérivationnel* a été très peu explicité dans PTQ. Trois remarques « en passant » – c'est-à-dire non présentées

explicitement en tant que telles et déjà présentées en citations ci-dessus – donnent une idée de comment ces arbres sont conçus. On a :

- D'après la Remarque i, (PTQ p. 255), il serait évident (anglais *obvious*) comment utiliser les arbres de dérivation en tant qu'expressions d'un langage non ambigu à la PTQ.
- D'après la Remarque ii (PTQ p. 263), un arbre de dérivation admet une seule traduction.
- D'après la Remarque iii (PTQ p. 255), une même expression graphique non élémentaire admet plusieurs arbres de dérivation se différenciant parfois de manière « authentique » (anglais *genuinely*).

On admettra ici la Remarque ii sans discussion : à chaque arbre de dérivation correspond donc une seule traduction. Le problème est donc de savoir si une expression graphique non élémentaire, du fait d'appartenir à plus d'un ensemble indexé d'expressions graphiques – c'est-à-dire à P_A et à $P_{A'}$, avec $A, A' \in Cat$ – sera associée à plus d'un arbre dérivationnel menant à des traductions différentes dans LI. Ou autrement dit, utilisant la terminologie de PTQ, le problème est de savoir si étant donné $expr \in P_A$ et $expr \in P_{A'}$, $expr$ sera associée à des arbres dérivationnels « authentiquement » différents.

On a deux réponses :

- Il est formellement impossible d'associer $expr$ à deux arbres dérivationnels différents à la PTQ façon Montague.
- Il est formellement possible d'associer $expr$ à deux arbres dérivationnels différents à la PTQ façon Chambreuil.

La différence entre les deux réponses découle directement de la différence dans la conception des arbres à la PTQ signalée ci-dessus : si l'on indexe les expressions – ce que PTQ ne fait pas – on ne perd pas l'ambiguïté d'assignation catégorielle et, du même coup, on obtient deux arbres dérivationnels « authentiquement » différents et par là la possibilité d'obtenir des traductions différentes. En revanche, si l'on n'indexe pas les expressions dans les nœuds, on perd l'ambiguïté et on n'obtient pas deux arbres dérivationnels différents (« authentiquement » ou pas) et on perd, par là-même, la possibilité d'obtenir des traductions différentes.

Que faire de la Remarque (i)? Si comme dans PTQ on n'indexe pas les expressions, alors la Remarque (i) devient crédible. On peut en effet imaginer comment passer des arbres dérivationnels à la PTQ au « langage non ambigu » à l'UG, à condition, bien entendu, d'éliminer des arbres dérivationnels tout ce qui vient de la démarche algorithmique utilisée¹². Si en revanche, on utilise des arbres avec des expressions indexées, la Remarque (i) ne peut pas subsister : l'absence d'ambiguïté dans un « langage non ambigu » à la UG est définie sur les opérations et non sur les expressions indexées et/ou les CBF.

¹² Et qui introduit la différence non « authentique ».

Dans l'analyse de UG, nous avons fait état de deux procédés pour obtenir l'assignation ambiguë des expressions graphiques non élémentaires. L'un, le procédé (ii) de la Section 2.1, a recours à des expressions graphiques élémentaires assignées de manière ambiguë à des ensembles indexés différemment. C'est le procédé qu'il devrait être possible d'utiliser – si ses répercussions systémiques n'étaient pas néfastes – pour rendre compte des exemples classiques du type *le juge, la belle ferme la voile*, etc. Or ce procédé (ii) n'est pas utilisable en PTQ en raison d'un fait déjà signalé (le non-sens formel qui advient). Il reste donc à PTQ la possibilité d'obtenir des arbres de dérivation différents pour une même expression graphique non élémentaire à une double condition : insérer dans les nœuds des arbres des expressions indexées et ne pas avoir incorporé aux expressions non élémentaires des expressions élémentaires appartenant à des ensembles indexés différemment.

4 Réponses aux questions sur le traitement de l'ambiguïté dans la Grammaire de Montague

Les langues naturelles sont naturellement ambiguës : voilà dit en termes simples ce qui est dit dans la formulation III de la Section 7.2 du document ci-avant EGM. Mais cette formulation III, exprimée en termes montaguiens moyennant les notions de « langage non ambigu » et de la relation *R*, n'explicite qu'une des facettes de l'ambiguïté des langues naturelles. Les formulations VI et VII en ajoutent deux autres. Ces formulations-ci expriment un fait têtue : à un moment donné et dans une circonstance donnée, une langue naturelle n'est pas le fruit d'une convention explicite qui, comme dans les langages logiques, repose sur l'assignation d'entités perceptivement différentes – en général, graphiquement différentes – à des ensembles dont l'appartenance est justement explicitée en pointant sur les différences graphiques. C'est le contraire qui arrive dans les langues naturelles : dans celles-ci, pas de convention explicite et surtout pas d'assignation détectable par des caractéristiques graphiques autres que l'énumération extensionnelle des unités assignées à des ensembles différents et pas de non-intersection entre ces ensembles. « It is a fact » comme dirait Popper, et ce fait, crucial, omniprésent, envahissant n'a pas été intégré à la problématique à traiter dans son système ni par Montague lui-même, ni par ceux qui ont présenté, analysé, diffusé la Grammaire de Montague *stricto sensu*.

Les formulations III, VI et VII de EGM Section 7.2 n'épuisent pas la problématique générale de l'ambiguïté dans les langues naturelles. Ni dans EGM ni dans cette Annexe on n'a évoqué le problème de la référence, ce qui, pour le faire sérieusement, suppose, encore une fois, de distinguer soigneusement UG (Section 4 « Semantics: Theory of reference »), l'univers de dénotation de LI, l'interprétation induite que l'on devrait trouver selon les définitions de UG et celle à laquelle on devrait arriver selon telle ou telle interprétation de PTQ. On trouverait sur ce chemin dangereux le problème des homographes, bien connu des lexicographes : le *cousin* membre de la famille et le *cousin* insecte, en français, la *pata* femelle du canard et la *pata* jambe d'un animal en portugais ou en espagnol, sont des exemples familiers d'une problématique, elle aussi constante, dont on ne voit pas de solution dans le formalisme des interprétations induites de UG, et pour lesquels aucune

solution n'est proposée ni dans EGM ni, à notre connaissance, ailleurs dans, bien entendu, le cadre de la Grammaire de Montague. L'homographie est une situation courante, elle aussi, et elle ne peut être envisagée sans envisager le système interprétatif de LI et son adéquation – ou non – à traiter des problèmes bien répertoriés des langues naturelles.

Dans la Section 1 de cette Annexe nous avons posé les deux questions suivantes, reproduites ici :

- Est-ce qu'en syntaxe on peut avoir des expressions graphiques, élémentaires ou non, associées à plus d'une catégorie?
- Si la réponse à la question précédente est oui, quelles sont les conséquences que cela entraîne pour les autres pièces du système de la Grammaire de Montague, en particulier pour la traduction en LI et/ou l'interprétation de la syntaxe dans un univers de dénotation?

Ces questions ont été explicitées et traitées dans ces cadres différents que sont UG et PTQ. On synthétise les réponses, ci-avant justifiées, auxquelles on a abouti :

- On peut formellement associer des expressions graphiques, élémentaires ou non, à plus d'une catégorie, et cela aussi bien dans UG que dans PTQ.
- Les répercussions systémiques sont telles que :
 - Dans UG, on aboutit à un non-sens formel aussi bien pour les expressions élémentaires que pour celles qui ne le sont pas, et ce non-sens formel ne peut pas être pallié par des stipulations ad hoc, car celles-ci introduisent des résultats inadéquats par rapport à l'observation de ce que l'on veut exprimer ;
 - Dans PTQ, on aboutit à un non-sens formel pour ce qui concerne les expressions graphiques élémentaires; pour ce qui concerne les expressions non élémentaires, si l'on donne aux arbres de dérivation des caractéristiques formelles (mal explicitées dans PTQ) telles que ces arbres aient le même pouvoir expressif que les expressions d'un « langage non ambigu » à la UG, alors on ne peut pas associer des arbres de dérivation « authentiquement » différents à des expressions assignées par des CBF à des ensembles d'expressions différemment indexés et par là, l'ambiguïté reste inexprimée au niveau de la traduction; si, au moyen de ces arbres, on indexe dans les nœuds les expressions, alors on peut exprimer, moyennant ces arbres, des ambiguïtés « authentiques » (mais ces expressions non élémentaires ne doivent pas être constituées d'expressions élémentaires assignées de manière ambiguë à des ensembles indexés différemment).

Ces conclusions justifient, confortent et élargissent l'Ob. 3 de la Section 3.2.1 du document ci-avant EGM. Ce document EGM est partiellement centré sur la problématique de l'ambiguïté d'assignation catégorielle des expressions élémentaires. On y a insisté et on insiste ici sur le fait que l'impossibilité d'exprimer dans la Grammaire de Montague de manière adéquate les ambiguïtés d'expressions élémentaires, moyennant leur attribution à des ensembles indexés différemment, ne veut pas dire impossibilité absolue de rendre compte de ce type d'ambiguïté. On a utilisé pour ce faire dans EGM l'artifice d'introduire des symboles « vides », c'est-à-dire des symboles sans contrepartie graphique, ce qui

oblige à avoir recours dans les arbres de dérivation à des opérations d'effacement, ce qui, à son tour, augmente de manière non contrôlée le pouvoir expressif de la syntaxe. On est ainsi en plein dans la problématique posée par l'existence des trois zones d'intervention où on peut traiter la naturalité de la langue naturelle (cf. EGM, Section 7.2 in fine) : on a le choix de compliquer la syntaxe et/ou de compliquer les Op-K permettant d'obtenir l'algèbre dérivée de LI, et/ou de compliquer les postulats de signification. Dans EGM le choix a été fait de compliquer la syntaxe, sans intervenir sur les autres zones. Le système d'ensemble est tellement peu contraint que d'autres choix auraient été possibles, y compris à l'intérieur de la syntaxe. Mais ce qu'il nous intéresse de signaler ici c'est que, quoiqu'on fasse, on aura besoin de stipuler un « truc » qui n'est pas nécessaire pour traiter un langage logique : la naturalité de la langue naturelle est implacablement à l'œuvre.